The dark side of the amazing phenomenon of muscle memory on social media.

Master Thesis for the degree of Human-Computer Interaction.

Maria Elisavet Kontogeorgou
m.e.kontogeorgou@students.uu.nl

Supervisors
Dr. Almila Akdag
Dr. Christof van Nimwegen

Utrecht University

Department of Science
Utrecht University, the Netherlands

December 28, 2022

Abstract

Mobile phones tend to be an extension of our hands, and the apps in them consume most of our time. Spending that much time on mobile apps and, more specifically, on social media apps results in building muscle memory to perform many interface-related actions. When a task is repeated, it becomes part of procedural memory. That type of memory dedicated to movement is called "muscle memory", which allows one to perform actions without seeing the screen and, often, without even thinking.

The mechanism of muscle memory minimises the load capacity for the users but, on the other hand, makes them vulnerable to the app creators of these apps. The ability to take advantage of users' muscle memory to promote the app's goals (new feature launch, click-on ads) came recently into the scientific research scope and still needs to be explored in depth. Instagram in many countries changed the tab bar menu and replaced the most used "create" action item with the new feature "Reels". The new feature, "Reels", contains short videos from random content creators. Instagram did that not for the users' benefit but to promote its new feature. By replacing the most commonly used functions with 'new' functions at the exact location, Instagram expects the users to press the button they actually did not intend to - which is exactly what a dark pattern is by definition.

In this research, through a user study, we evaluated their muscle memory using the tab bar menu on social media apps. The study was conducted remotely, targeting the Greek population. We did that because, in Greece, Instagram has not yet launched the "Reels" feature. We examined how that change affected the users' experience and how the users perceived the change. We found that higher usage frequency builds stronger muscle memory. Moreover, the Greek users needed more time and had more miss clicks on their attempts to perform the "create" task on the interface containing the Reels feature, compared with the initial interface. Combining the previous results with users' perceived feelings lead us to conclude that the tactic of taking advantage of users' muscle memory from Instagram to promote a new feature can be characterised as a dark pattern.

Lastly, we introduced a new term, "collective muscle memory", a phenomenon where users apply muscle memory built from one social media to another. We examined that term by testing whether Instagram users can benefit from their muscle memory to perform similar tasks on Pinterest. To do so, we examined Instagram users without prior experience with Pinterest. We found that frequent Instagram users have statistically significantly lower execution time on Pinterest tasks than the less frequent users with no related muscle memory. Pointing out that "collective muscle memory" is a valid term in the domain of social media apps.

Keywords: muscle memory, dark patterns, social media, collective muscle memory

1 Introduction

Social media are ubiquitous in society today. Billions worldwide use social media to share information and make connections [1]. More than half of the world now uses social media (58.4%). Specifically, 4.62 billion people around the world use social media, and 424 million new users have come online within the last 12 months. Interestingly, the average daily time spent using social media is 2h 27m [18].

1.1 Social Media & Dark Patterns

Understandably, social media apps are the gold rush of our time. The companies running them are implementing design patterns to increase their user base and usage time aiming to increase their market size. Sometimes these design patterns could be classified as dark patterns.

Dark patterns are when an interface is maliciously crafted, with a solid understanding of human psychology, to deceive users into performing actions they did not intend to do [5]. Most of the already researched dark patterns on social media have to do with data privacy [36]. From the Privacy "Zuckering" [1], to "Roach Motel" [2] technique and misdirection; all of them aim to make users accept or not be able to decline or edit preferences about the collection/storing/editing of their personal data [4, 29]. The European Data Protection Board (EDPB) adopted guidelines on dark patterns in social media platform interfaces. The guidelines offer practical recommendations to designers and users of social media platforms on how to assess and avoid so-called "dark patterns" in social media interfaces that infringe on General Data Protection Regulation (GDPR) requirements [11].

Another popular dark pattern on social media apps is the so-called "Fake Notifications". Where these apps send notifications without an important reason, like "A person that you are following made a post after a long time" or "People with similar interests are following the X person, follow him too", making the users check their notifications and eventually use their social media apps more regular [6].

The disguised ads are a dark pattern with almost the same popularity among social media apps. As all social media apps are free of charge for their users, their revenue is based on ads. Most of them are well fitted between the content and, most of the time, look quite similar, making it extremely difficult for the users to recognise them [25].

[1] This dark pattern was named after Facebook CEO Mark Zuckerburg and referred to Facebook's early privacy settings, which in general made it difficult for users to find and, as a result, making it easy for users to 'overshare' information, including profile photos, telephone numbers and more.

[2] The roach motel design makes it incredibly easy for a user to get into a certain situation but difficult to leave. A popular case is the deletion of an Amazon account, which is the only way to erase a user's purchase history and collected data. On the one hand, it is painless to sign-up for an Amazon account but on the other hand, deleting the account is a bit of an Easter egg hunt and not an enjoyable one.

In this research, we explore one case which has not been researched a lot and has to do with the intention of social media apps to promote features taking advantage of their users' muscle memory. Yildirim, in his thesis, examined this case using the latest example of Instagram, where they have launched a new feature, "Reels", and promoted it taking advantage of users' muscle memory [38]. One of the weak points in his study was that he used eye tracking in a desktop lab environment. An environment that is not the one in that users have built muscle memory. By that, we mean that Instagram users use Instagram as an application on mobile devices, so any muscle memory is attached to the thumb movement and not to a mouse movement. That would also be one of the key differences with our research, where we will conduct an experiment on each participant's mobile device.

1.2 Muscle Memory

When humans need to perform a task or solve a problem, they usually take the path of minimum effort, not because they are lazy, but because they attempt to be efficient and satisfy their many goals as fast and easily as possible. To do so, they subconsciously match the current situation to past encounters. If, in a similar situation, a specific action has been successful in the past, that action will be selected to be applied again despite the alternatives.

When a task is repeated a lot, it becomes part of the procedural memory: it is almost as if we have a dedicated "muscle" that deals with it. Not surprisingly, a type of procedural memory dedicated to movement is called "muscle memory." This muscle memory allows us to unlock our phone using a diagram, type our PIN on an ATM, or log in to our computer; we can complete these tasks on autopilot [15].

1.2.1 User Interfaces

Consistent placement of action items (like "save", "back", and "home" buttons) in a user interface not only help users learn the system but also help their muscles "learn" the system as well. A user's constant and continual use of a product's user interface (UI) causes their muscles to "memorise" certain movements and actions. For example, when a user's muscle memory instructs them that the "Create" button is always the middle action button on their tab bar menu, the user will eventually navigate to that spot through muscle memory without requiring any visual confirmation that the "Create" button is actually there. This is why consistency in graphical placement within a user interface system is so important.

1.2.2 Collective Muscle Memory

As previously mentioned, users rely on their muscle memory to perform most of their frequent actions in their favourite applications. On top of that, it has been noticed that users are taking advantage of their muscle memory of every other similar website/application they have ever encountered to know how to interact with the current website or application. Further, the result of that action is assumed to be exactly the same as the results achieved in the past, turning the past experiences into current expectations [15]. In this research, this is what we define as collective muscle memory.

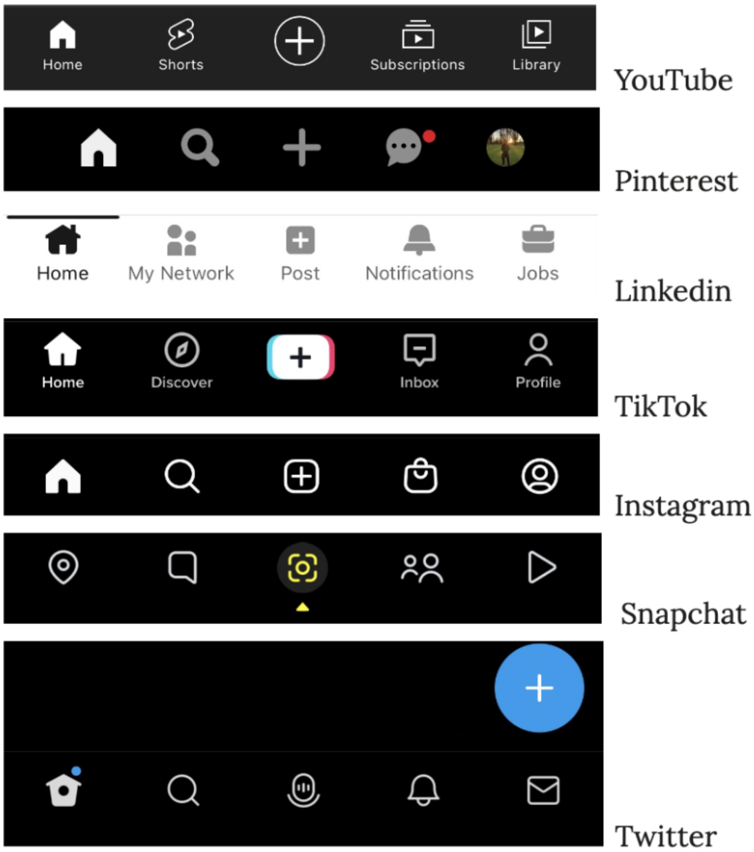

Figure 1: Example of Social Media Tab Menu

1.2.3 Collective Muscle Memory On Social Media Applications

Most top social media apps use similar design patterns on their interfaces. The most characteristic one is on the tab bar menu, where almost all of them are using the same action items at the same places, "home", "search", "create", "inbox", "notifications", "profile" (Figure 1). As a result, we expect that users interacting with more than one social media app will be able to utilise the muscle memory created by one app to another one, contributing to the creation of collective muscle memory. So, users' muscle memory is applied within and between apps. Meaning that they can take benefit of the muscle memory within the app by executing tasks with mechanical movements, and also between similar apps in executing tasks in similar environments even without having prior experience with these applications.

Having said that, we derive the assumption that users who frequently use at least one social media app do have a muscle memory dedicated to social media apps in general. Furthermore, any change breaking that norm is at least annoying for the users [17]. Snapchat faced that when they made a significant change to their interface [13]. Instagram, for example, in many European countries, did change their tab bar menu and replaced the probably most used "create" action item with the new feature "Reels" [16]. The new feature "Reels" contains short entertainment videos from random content creators or Instagram users. Instagram did that not for the benefit (convenience) of the users but instead in order to promote its new feature. By replacing the most commonly used functions with 'new' functions at the exact location, Instagram expects the users to press the button they actually did not intend to - which is precisely what a dark pattern is by definition.

1.3 Research Problem & Questions

In order to set up the basis of this research, we formulated the following research questions and sub-research questions:

RQ1: Could Muscle Memory usage from Social Media Apps turn into a Dark Pattern?

SubRQ1: Does higher Usage Frequency of Social Media Apps develop stronger Muscle Memory?

RQ2: Do the users perceive the Social Media Apps' technique of using Muscle Memory to promote New Features as a Dark Pattern?

RQ3: Can we apply Muscle Memory acquired from one Social Media App to another, proving the existence of "Collective Muscle Memory"?

1.4 Outline

In the following section, we present the results of a more targeted literature review aiming at the domain of the muscle memory effect on user interfaces (UI) and the dark patterns awareness from the users. Afterwards, we describe the study's research methodology and explain the experiment design and the tools we used. In Section 4, we present the analysis together with the results. We conclude this paper by providing the key findings of this project along with the main limitations and proposals for future work.

2 Literature Review

Initially, we performed a literature review to explore the domain of the dark pattern on social media applications and the usage of muscle memory on interfaces. The findings have been presented in the introduction part. Following that initial literature review, we proceed with a more in-depth review to understand how the muscle memory effects can be tested and measured and how we can assess users' awareness of dark patterns. In the following sections, we present our findings.

2.1 Measuring The Muscle Memory Effect On User Interfaces

Muscle memory is a term commonly used for the sort of embodied implicit memory that unconsciously helps us perform various motor tasks we have somehow learned through habituation, either through explicit, intentional training or simply as the result of unintentional or unconscious learning from repeated prior experiences. In scientific terminology, such memory is often designated as "procedural memory" or "motor memory" because it enables us to perform various motor procedures or skills intuitively. We do these actions without thinking of how the procedure should be followed and how to proceed from step to step. Some examples of such muscle-memory motor skills of performance are walking, swimming, riding a bicycle, tying one's shoes, playing the piano, driving a car, or typing on a keyboard [30].

Most muscle memory papers refer to the actual memory of muscles related to physiotherapy, rehab, etc. In some papers, there was a separation from that definition, and they proposed alternatives like memorability, motor learning, and implicit memory. We used these terms to extend the keyword used in our literature review [14].

The first research that refers to the effect of muscle memory on the user interface and, more specifically, on a UI's menu component has been public since 1988. Wadlow Maria G et al., in their guidelines for the Andrew System interface, have explicitly defined muscle memory as the ability to make a motor movement without conscious attention. They also point out that the menu UI components take advantage of muscle memory by positioning options in the same places on different menus [35].

In our quest for more recent research on the specific domain, we found some work that either had to do with the assessment of the muscle memory effect on memorising passwords or on the eyes-free handling of touchpads. In all cases, the success criteria of the muscle memory evaluation were the time (how fast users performed the tasks) and accuracy (how many correct attempts they had - error rate). In their research, Chong Ming Ki and Marsden Gary examined the advantages of muscle memory as a way to memorise gestures used for authentication in mobile phones. Their main tactic was to build and then test the muscle memory after some rehearsal sessions. Their success criterion was the number of successful PIN gesture retrievals (accuracy) [7].

Their subjects were given a 24-hour rehearsal period to memorise the passwords. Six days later, the researchers requested the subjects to recall their given passwords. The results were not as good as traditional authentication methods (numeric PIN). The researchers suspect these low accuracy scores to have been caused by the subjects not having enough time to practise their passwords. One of the subjects had explicitly mentioned that she needed more time to memorise her given passwords.

In a similar context, Van Koningsbruggen, Hengeveld, and Alexander examined the usage of muscle memory in recalling passwords; this time using embodied passwords. Again the success criterion was the number of successful attempts to recall the password. For the experiment, the users needed to build muscle memory by practising, and they were examined after a day and then again after a week [34]. All participants could recall and reproduce their embodied passwords during the last session. One participant forgot all alphanumeric passwords, and another participant forgot one of the alphanumeric passwords. Based on this, people can recall embodied passwords, and the preferred learning style does not influence their ability to recall a particular type of password.

In our case, muscle memory will already be built as our users will have spent time in everyday life in the examined environments (social media apps).

As previously mentioned, another perspective on researching the muscle memory effect involves assessing the eyes-free usage of touchpad devices. Lu et al., explored the possibility of eyes-free text entry using one thumb where the users do not look at the touchpad but rely on their muscle memory only to move the thumb and tap. Participants were instructed to input phrases using the touchpad "as fast and accurately as possible" [21]. After practising only a few words, the results showed that users could transfer their typing skills from normal conditions to eyes-free use.

A related study Crossland et al., evaluated the impact of muscle memory on interaction mechanisms with in-vehicle touch screens. To do so, they displayed a single, white square button (12 x 12 cm) at a time on the vehicle touch screen and asked participants to press the button as quickly and accurately as possible while driving. Task time was significantly higher for muscle memory than for peripheral vision conditions. Data showed that participants made more errors – did not press the button correctly the first time – during muscle memory conditions compared to peripheral and foveal vision conditions. However, their primary task was the driving task, which could have affected the outcome [9].

2.2 Dark Patterns Awareness

Another important aspect of our research is to explore the available literature on assessing dark pattern awareness. In simpler words, to understand if and to what extent users recognise dark patterns in UI and how that affects them. Again, there needs to be more research for the specific subdomain of our research. Moreover, there needs to be more literature on analysing users' experience and acceptance of dark patterns using quantitative methods. Most researchers use qualitative methods such as open questions or focus groups to collect the insights and the required data.

User interfaces containing dark patterns are also called "deceptive interfaces". Láng and Pudane, in their research on deceptive interfaces [19], performed a case study on Amazon's account deletion navigation and its effects on user experience. The dark pattern Amazon uses is called "the roach motel," so named because, like the insect's trap, users easily create their account, but they cannot delete it. It is an example of an unethical design practice because the right thing to do would be to make it as easy to cancel a subscription as it is to start one. Using open-end questions, they concluded that frustration was the most common word to describe the experience with Amazon's account deletion process; discouragement was also mentioned multiple times. A few participants touched on the idea of deception by expressing that they felt tricked, even before being further informed on the research topic. Lastly, all participants labelled their attitude towards Amazon as negative[19].

In his qualitative study [24], Maximilian aimed to find out how the end user perceives, experiences and responds to dark patterns. He had one focus group with 6 participants and 5 semi-structured interviews. After some initial questions, the participants were presented with a quick overview of dark patterns on the web and mobile environments and a brief definition of them. The perceived feeling of everybody participating was annoyance, although the reaction always depended on the experienced damage. For example, while the experience of facing behaviour-influencing elements made some participants want to quit their session on the application, it seemed that if the benefits of using the service/application outweighed the encountered negative aspects, users would continue to use the service. It was also pointed out that if users are dependent on a service (like social media apps), they are inclined to accept possible minor negative aspects [24].

The first attempt at quantitative research regarding the perception of the dark patterns in UX and the reaction to them has been made by Di Geronimo et al., [10]. They analyzed dark patterns in 240 popular mobile apps and conducted an online experiment with more than 580 users in order to understand how the effect of dark patterns in such apps is perceived. In that regard, they used the following two questions:

- Is this user interface crafted to trick the users into doing things they do not want to do or try to manipulate the user in some way? 'Yes/No/Not sure.'

- How annoying was each malicious design? Likert Scale from 'Very annoying' to 'Not annoying at all'.

The majority of the participants were either not able to detect dark patterns or were not sure about them. That was justified by some participants' explanations that dark patterns are so widely spotted in modern applications that they become part of the expected interaction flow. Nevertheless, the researchers did not refer to the results of the second question regarding the perceived annoyance from the malicious designs [10].

A second and more completed quantitative research was performed by Bergman [2] on her master thesis focusing on the deceptive UIs from users' perspective. Her main goal was to measure the severity -or 'darkness'- of the various types of dark patterns, turning that eventually into a 'Dark Pattern Darkness Score' (DPDS). Furthermore, she developed a 'System Darkness Scale' (SDS), which can be used to evaluate the 'darkness' of a system as a whole. She designed two versions of an e-commerce website, with and without dark patterns so that participants could interact with it. The participants were randomly assigned to either the A or B version of the experiment and completed various shopping tasks. After that, they had to evaluate their experience in either the "Dark" environment or the "Bright" one.

On her SDS scale, she used the following items:

1. The system tricked me into performing certain actions that I did not intend to do.
2. The system performed certain actions I was not aware of.
3. The system performed certain actions without my consent.
4. The system pushed me into spending more money than I originally anticipated.
5. I felt deceived/misled by the system.

The results showed no significant difference in User Experience between the two versions. That has been justified by the wide usage of these patterns in daily life, having as results to be felt as part of the process by the users.

Another factor worth mentioning is the one of "dark patterns blindness". Di Geronimo et al. [10] introduced the term to explain why most users cannot recognise dark patterns in mobile applications. However, in his study, when the participants were informed of the potential presence of dark patterns in their experiment, they became more capable of spotting them. Luguri and Strahilevitz [22] showed that mild (i.e., more subtle) dark patterns become more easily unnoticed than aggressive ones. Furthermore, M. Bhoot, A. Shinde, and P. Mishra [23] found that the ability to identify a dark pattern is correlated with its frequency of occurrence and the frustration it provokes.

Summarising our literature review, we conclude that the key metrics to measure muscle memory are time (how fast a user performs a task) and accuracy (how accurately a user performs a task, usually counting the number of attempts or missed clicks). On top of that, a common phrase that the researchers used to trigger the user's muscle memory was "Perform the X task as fast and as accurately as possible". We used both the metrics and the phrase in our experiment too. Moreover, as it might be difficult for the participants to identify dark patterns, we will evaluate the participants' perceived feelings. To do so, a selection of items from the SDS score combined with the items from the study of Di Geronimo et al. [10], is used to measure the dark pattern awareness from the users.

3 Research Methodology

In this section, we present our conceptual framework and the hypothesis we set to answer our research questions. We analyse our experiment design, describing every step of the experiment process that the users had to follow together with our collected data points. After that, we describe our target population and the reasons behind its selection, and we close with a presentation of possible tools for that kind of experiment, focusing more on our selected one.

3.1 Conceptual Framework

Based on our literature review findings, we formulate the conceptual framework depicted in Figure 2. Literature shows that having many repetitions of the same action results in building muscle memory. We will test if that theory applies to the social media apps domain (H1).

We have also found that social media apps, specifically Instagram, use users' muscle memory to promote new features. They do that by replacing items that users have used to click to perform a specific action with a different item. They do that expecting their users to click on them accidentally, and by doing that, they will be introduced to the new feature. We will test if that tactic can be characterised as a dark pattern (H2). Finally, by analysing the interface of the most popular social domain apps, we found many similarities in the design patterns. The most prominent is the tab bar menu (Figure 1). We assumed that muscle memory built from one social media app could be used in a different one, even without prior usage (H3).

Instagram in many European countries changed the tab bar menu and replaced the probably most used "create" action item with the new feature "Reels". The new feature "Reels" contains short entertainment videos from random content creators or Instagram users. By replacing the most commonly used functions with 'new' functions at the exact location, Instagram expects the users to press the button they actually did not intend to - which is exactly what a dark pattern is by definition.

In this thesis, we simulated the release of the "Reels" feature to the Greek population. We did that because, in Greece, Instagram has not yet launched the "Reels" feature that we described previously. So we examined how a change will affect the users' experience. If it will affect the efficiency (time to perform tasks) or their accuracy (number of correct

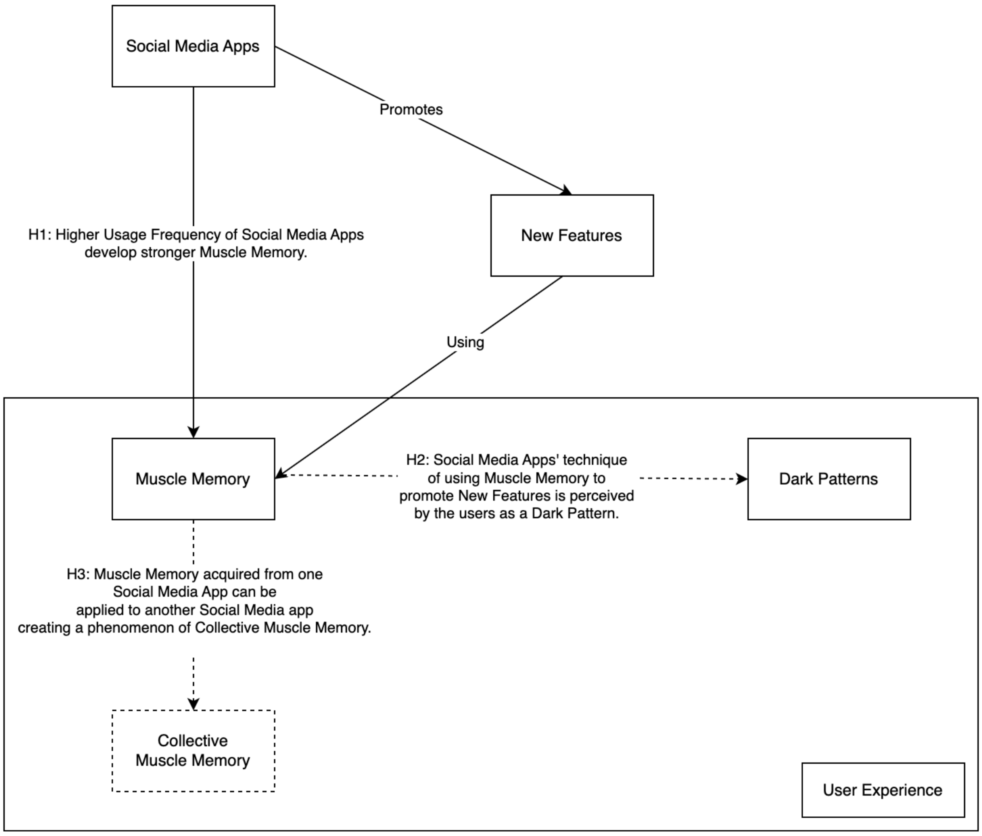

Figure 2: Conceptual Framework

clicks to perform tasks). On top of that, we analysed how the users perceive this change.

3.2 Hypothesis Set

To be able to respond to the research questions that we presented earlier, we formulated the following hypothesis:

- H1: Higher Usage Frequency of Social Media Apps develop stronger Muscle Memory.
  - H1. A. Higher Instagram usage frequency will be translated to faster execution of the tasks on the Greek interface known to them (efficiency).
  - H1. B. Frequent Instagram users will have more accidental clicks on the Dutch interface, new to them, than the rest (accuracy).
  - H1. C. The users who already use the Dutch interface (dutch population) will be faster and more accurate due to muscle memory than those who use the Greek interface (baseline).
- H2: Social Media Apps' technique of using Muscle Memory to promote New Features is perceived by the users as a Dark Pattern.

- H3: Muscle Memory acquired from one Social Media App can be applied to another Social Media App creating a phenomenon of Collective Muscle Memory.

3.3 Experiment Design

To validate our hypothesis, we performed an experiment. We examine the muscle memory of social media app users and, more specifically, the ones from Instagram. As we previously mentioned, Instagram two years ago in many European countries changed their tab bar menu and replaced the probably most used “create” action item with the new feature “Reels” [16].

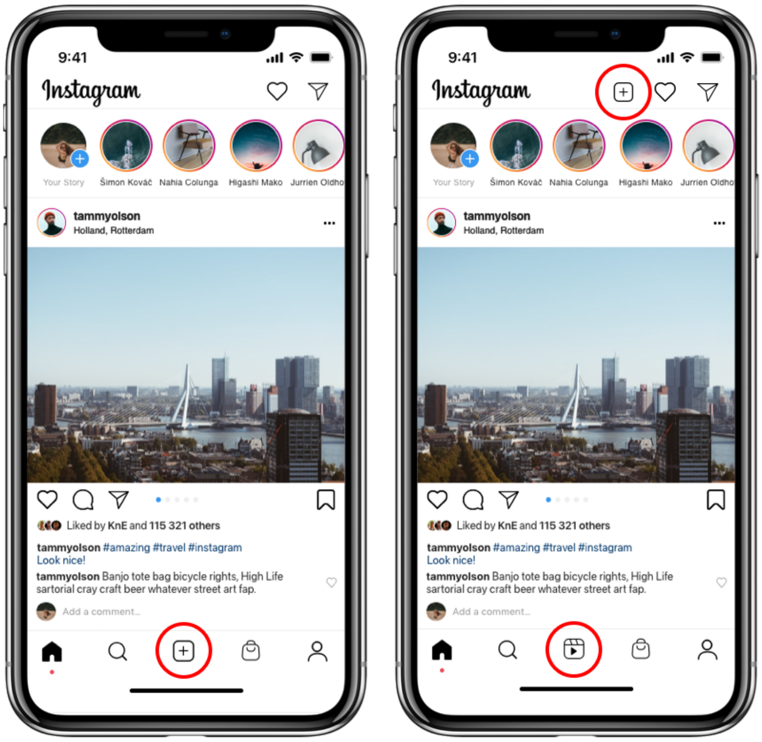

Figure 3: Left: Greek Interface - Right: Dutch Interface

With the “create” action item, the users had the ability to share an image or a short video through the creation of a post or a story (stories are creations that usually are publicly available for 24 hours after the creation). The “create” action button was located in the tab bar menu’s prominent position, the centre. Instagram, instead of adding a sixth action item for the Reels, replaced the “create” action item with the “Reels” and moved the “create” to the top right corner. In Figure 3, the two UIs, with and without the “Reels” feature, are depicted. In this way, Instagram took advantage of the existing muscle memory of its users. Meaning that while they would try to create a post or a story by clicking on the usual spot in the tab bar menu, they would end up in the new “Reels” feature screen.

Below a description of the experiment can be found: We present the experiment flow in Figure 4. We started by asking for demographic data, then moved to the questions aiming to quantify their Social media usage. After that, the users started to perform the tasks of the experiment, which are explained in more detail in the following paragraphs. In the end, there was an open question to collect any additional feedback. There was no elimination of participants as we were looking for all types of social media usage, even the ones that have never used Social media apps before.

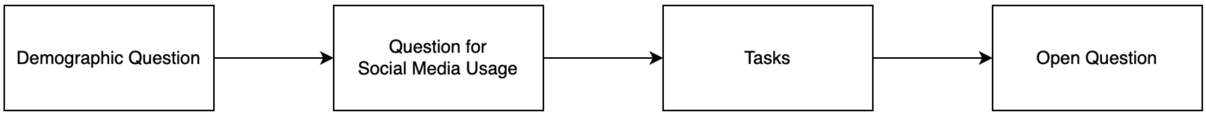

Figure 4: Experiment Flow

Initially, we evaluated the social media apps usage **frequency** of all the participants for a selection of social media apps. To do so, we used the following 7-point response scale question for Social Media Usage (SMU) as it has been used by Boer et al. [3] in their research.

- How many times a **day** do you check the following social media apps?

We selected the most used social media apps [18] that follow the same design pattern on their tab bar menu Figure 1. By having this question, we could understand the type of users we had in our participant pool, from rare social media app users to heavy ones.

Right after, we asked them to perform a familiarisation **task** on the **Original Interface** of Instagram (Greek version) in order to understand the concept of the experiment and get used to handling the mock-up. That familiarisation task asked them to go to their profile and look for a specific post as quickly as possible. The intention of this task was for them to click on the tab bar menu using their muscle memory and then look for a post. On this task, although we gathered time (efficiency) and accuracy data (accuracy), we did not use them in our analysis.

Since they were done with the familiarisation task, we asked them to perform a second **Task** on the **Original Interface** of Instagram (Greek version), and we measured their efficiency and accuracy. The task was to go to the search screen and look for a specific post. To do so, the user had first to click on the search button on the tab bar menu. We deliberately asked to focus on a goal other than clicking on the tab bar menu, although that was what we were measuring. We did that to avoid any bias that might appear from users that by understanding the goal of the experiment, they would change their usual behaviour within the application. We wanted them to believe that we were looking for something else and let their muscle memory work. We measured the time of execution and how fast they clicked on the search button on the tab bar menu. On top of that, we measured how accurate they were, meaning how many clicks they did to complete the task. As the search button is placed on the tab bar menu, we expected that participants would need only one click for this action. These metrics have been used to examine part of the first hypothesis (H1.A).

Furthermore, we asked them to perform a third **Task** on an **Altered Interface** of Instagram (simulating the version that exists now in many European countries, including The Netherlands) and measure their efficiency and accuracy. The task was to click on the create button, which on this interface has been moved to the top right corner of the application. We expected they would accidentally click on the location where the create (+) button used to be, and now Instagram has placed the new Reels button. That is why we also captured the miss clicks, especially those in the position where the (+) button used to be. Like the previous task, we measured their time of execution and how fast they clicked on the create button in the top right corner. On top of that, we measured how accurate they were, meaning how many clicks they did to complete the task. As the search button is placed on the initial screen presented to them, we expected that participants would need only one click for this action. These metrics have been used to examine the rest of the first hypothesis (H1.B,C).

After the third task, we asked them to fill in a questionnaire to describe their experience with the new design evaluating their perceived feelings. These were used to evaluate the dark patterns awareness. The questions were chosen from the questionnaire proposed by [2] and adjusted to fit the specific case. Bergman in her thesis, she proposed five items but not all matched with the dark pattern we were examining. As a result, we excluded the items related to giving consent by the users and spending money and kept only the following related ones.

- Instagram tricked me into performing certain actions that I did not intend to do. [Strongly Disagree - Strongly Agree]
- I felt deceived/misled by Instagram. [Strongly Disagree - Strongly Agree]

These questions were used to validate the third hypothesis.

Due to time and resource availability, we took the opportunity to examine one popular dark pattern used by social media apps, the disguised ads one. As has been mentioned, disguised ads are designed to look almost identical to the actual content so that the user will be misled and unintentionally interact with them. So in this task, we asked the participants to look for a post with a particular number of likes. We did that as we wanted the participants to look at all the posts carefully and in an area where the ads used to have the only differences with the actual posts. As it can be seen on Figure 5 Instagram designed the sponsored content to look the same as an original post adding one call to action container (buy, sign-up, download) between the image and the standard action items (like, comment, share). So the actual goal was to have them look closely at more than 4 disguised ads while looking for the requested post. After that, we asked the following questions to evaluate how the participants anticipated the ”disguised ads” dark pattern on Instagram.

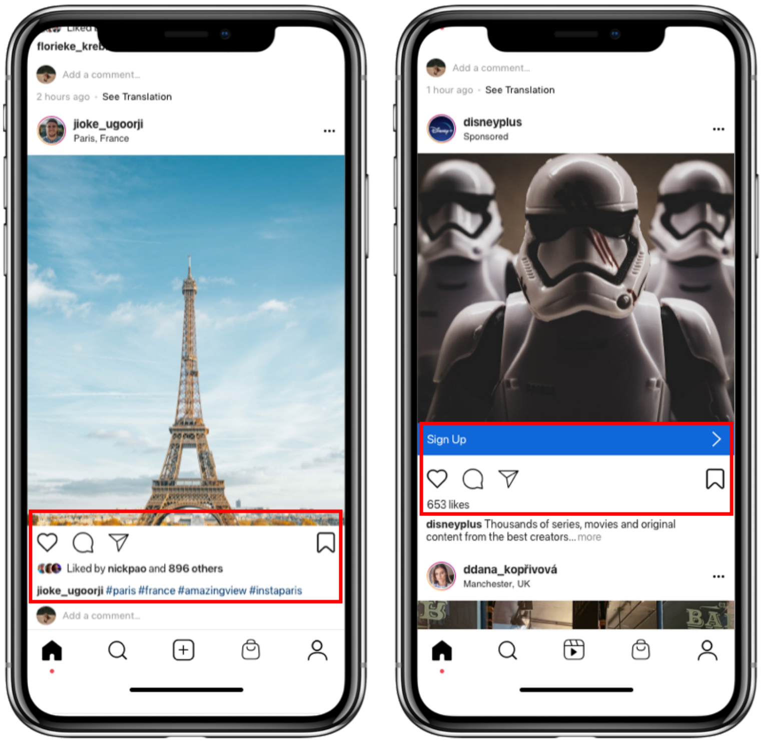

Figure 5: Instagram Post vs Disguised Ads

- Did you notice any sponsored content? [No - Yes]
- How do you feel about this? [I do not mind - I find it deeply annoying]
- How frequently have you experienced such a situation? [Never - Very often]
- Instagram performed certain actions I was not aware of. [Strongly Disagree - Strongly Agree]
- Instagram performed certain actions without my consent. [Strongly Disagree - Strongly Agree]

All previous tasks examined all the hypotheses about building muscle memory and dark patterns. On top of that, we tested the effect of dark patterns related to muscle memory on the users. To test the last hypothesis, we asked the participants to perform two last tasks, similar to the Instagram ones, on a mock-up of Pinterest.

The reason we selected Pinterest among the other social media apps we presented in Figure 1 is that we wanted to test if the muscle memory of one social media (a popular one), like Instagram, can be transferred/applied in another social media (not that popular where no prior muscle memory exists). Based on the usage data, Pinterest is the least used social media app among the ones with the most similarity on their tab bar menus [18]. To have an analogy of the usage, almost 42% of the Greek population is using Instagram but only 16% Pinterest and TikTok 27%. So we asked the participants to perform a familiarisation **task** again on the mock-up of Pinterest. On this task, although we kept time and accuracy data, we did not use them in our analysis. For the last task, we asked them to go to the search screen of Pinterest and search for an image. Again our goal was to check how fast and accurately they would tap on the search button on the tap bar menu rather than the task of finding the image. Finally, we asked them for feedback and thanked them for their participation. All the screens from the experiment can be found in Appendix subsection A.2.

3.4 Participants

We targeted the Greek population living in Greece, where Instagram still has a previous version than the one in other European countries. In Greece, Instagram still uses the previous version with the "create" action item in the middle of the tab bar menu. So we planned to simulate the launch of the "Reels" feature to the Greek audience and examine their reaction to it. By that, we meant to compare their efficiency and accuracy in implementing tasks on the old and altered interfaces. On top of that, we evaluated how they felt about the sudden introduction of the new feature in a place they were used to having another functionality. Based on that outcome, we wanted to determine if muscle memory plays a role in social media apps and if users perceive the manipulation of that as a dark pattern.

Through convenience sampling, we reached participants that are Greek Instagram users living in Greece and having different frequency usage. Nevertheless, we aimed to recruit various age groups, which was not easy to achieve.

3.5 Materials & Design

In this section, we will present the tools we selected for the app's prototyping and for the data collection (measurement) during the experiment sessions from the users. We also displayed the designs from the app we created for the experiment.

3.5.1 Prototyping Tool

There are a plethora of UI design tools on the market, and the most popular are Figma [12], Sketch [31] and Adobe XD [8]. In this study, we will use Sketch to design the prototypes. Sketch, is the all-in-one platform for digital design, with collaborative design tools, prototyping and developer handoff. Sketch, is a digital design app for Mac and uses it for UI, mobile, web and even icon design, as it is a vector-based app. The main reason why one of our choices is Sketch is that we already have a licence in Sketch and have working experience using this tool [31].

3.5.2 Measurement Tool

For our experiment, we would like to measure users' task execution time, collection of completed tasks, and incorrect attempts. Ideally, we would also like to have either eye-tracking or heatmap-recording of our users performing the tasks on our prototypes to get some qualitative aspects of their attempts. As we were researching muscle memory and, even more specifically, the muscle memory built by using the tab bar menu on mobile, we believed that heatmap is more important than eye tracking. That is because the thumb works automatically; taking advantage of muscle memory, it should not use any eye movement. On top of that, eye-tracking for mobile devices is not that advance yet, as it requires a fixed position for one of the two items (recording device or recorded subject). Even small body movements such as head re-positioning interfere with eye tracking, rendering its results unstable and unreliable as soon as the head is not in a fixed position anymore [20].

Based on our research for the best measurement tool for our experiment, we found many different tools that offer some of the features we wanted to use. The final list of tools that covered our needs for the experiment was Useberry [32], Maze [26], UXtweak [33], Optimal Workshop [37], and RealEye [28]. All of them work with prototypes from Sketch [31], and Figma [12]. From that final list, the best solution was a measurement tool which could combine time tracking and heat mapping with a low-pricing model due to the budget restrictions of this project. Maze and Useberry were the best options for us, and we initially selected the first following some recommendations from UX industry ex-colleagues and friends. However, we discovered that the Maze application had issues with supporting prototypes with fixed-positioned elements. In our case, we wanted to have the tap bar menu fixed at the bottom of the screen in order to have first an as much as realistic experience as possible for the participants but, most importantly, to be able to examine the users' muscle memory on the tab bar menu.

Useberry is a user-testing platform that turns prototypes into actionable measures with the help of real-world testing. There are two major objectives that it achieves, testing and reporting. Testers are various real-world users that provide instant feedback on the prototypes. Reports are then generated based on that feedback to gather data. Both steps are essential in getting validation on the design in question. Creating these user tests is extremely simple, and it can be automated. Useberry offers 10 responses per month for free, so we had to upgrade to their basic plan giving us the ability to have up to 100 participants.

3.5.3 Designs

As mentioned earlier, we used the Sketch to design our mock-ups. Although we had a license, it was necessary to acquire a cloud subscription-type license to connect it with the tracking tools like Useberry. We designed over 25 screens for three fully interactive prototypes of Greek & Dutch Instagram and Pinterest. All the images that were used were from unsplash.com and the icons from thenounproject.com, both platforms offer irrevocable, nonexclusive, worldwide copyright license to download, copy, modify, distribute, perform, and use for free.

Some of the Instagram prototypes (Greek interface) can be found in Figure 6, Figure 7, Figure 8, and Figure 9.

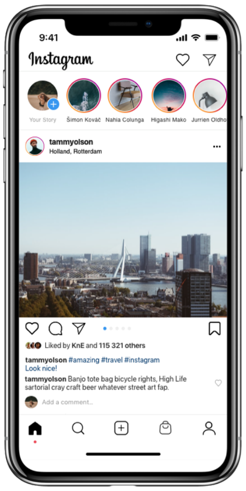

Figure 6: Home Screen

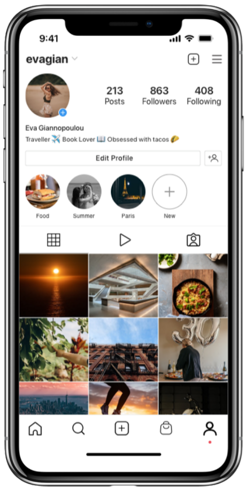

Figure 7: Profile Screen

Figure 8: Post Screen

Figure 9: Create Post Screen

Some of the Pinterest prototypes can be found in Figure 10, Figure 11, Figure 12, and Figure 13

3.6 Procedure

We run our experiment for almost two months (Sep 2022 - Oct 2022) remotely (online) through our network of friends, family and ex-colleagues in Greece. We had 87 participants, meaning that 87 unique users clicked on the link we shared and started the experiment. We had no supervision of them during the time of the experiment. We have just shared the link with them, and it was up to them where and when they would start the experiment. Of them, 60 completed the experiment, and the rest dropped at some point during the experiment. That was probably due to either not being willing to complete the experiment or accidentally closing the experiment by refreshing or closing the session. The only limitation we had was on the device, meaning that the link was functioning only on mobile devices (not on the desktop) as we wanted the examine the

Figure 10: Home Screen

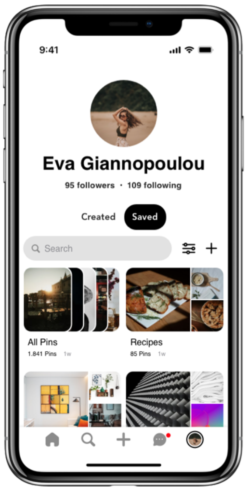

Figure 11: Profile Screen

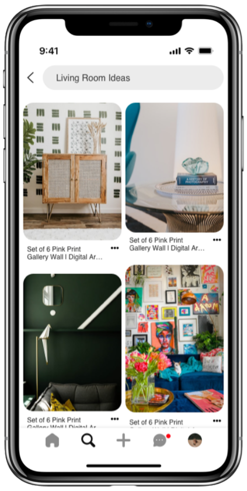

Figure 12: Search Screen

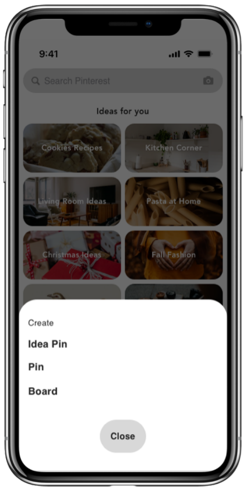

Figure 13: Create Post Screen

participant's muscle memory on their actual mobile device. On top of those, we run the experiment on a small sample of Dutch participants (9). These participants have been using the altered interface of Instagram (the one with the Reels feature) for more than a year and have built muscle memory to set the baseline for Task 3. We did that because we wanted to understand if the change in the interface had affected the time and accuracy of tapping the create (+) button for every user, no matter if they had developed muscle memory.

From Useberry, we were able to collect the following metrics:

- Operating System: Android, iOS
- Browser: Safari, Chrome, Opera
- Screen Recordings: Full-Screen recordings during the experiments for qualitative insights
- task time: The time in milliseconds it took the testers to complete each task.
- Miss clicks: Number of miss clicks
- Heatmap: For the qualitative exploration of testers' clicks/taps.

On top of that, we had two questionnaires to assess the usage frequency of social media apps and the dark pattern experience. We performed the following quantitative and qualitative statistical analyses based on these data.

4 Results

In this section, we will present our participants' demographic analysis and the results regarding the examined hypothesis.

4.1 Demographic Analysis

As we said before, 60 participants from the Greek sample completed the experiment in full, although five were excluded based on the outliers analysis we performed. The main reasons were extreme task completion time or a disproportionate number of missed clicks. As we saw from the recordings, these participants were facing either technical issues or were not focusing on the experiment (long time of inactivity and random clicks).

So our Greek sample population has been finalised to 55 participants who were mainly aged between 26 and 41 years old, probably due to the collection method, friends and family. More specifically, 83.6% (46) of the participants were aged between 26 and 41, 2 were older than 41, and 7 were younger than 26. More than half of them (56.4%) were using an Android device during the experiment, and the rest were using an iOS device.

The main browser used by the participants was Google Chrome (41.8%), followed by Safari (29.1%) and Facebook browser (16.4%); the rest of the participants (12.8%) used other browsers (Edge, Opera, Device native browsers).

We identified the following five profiles of usage based on the times they check a social media app per day.

- Rare Users — Less than 1
- Less Frequent Users — Less than 8
- Frequent Users — Less than 16
- High Frequent Users — Less than 24
- Heavy Users — More than 24

Based on these profiles, the most heavily used social media app was Instagram, with more than 55% of our participants checking it at least eight times per day, followed by Youtube (23%), Linkedin (11%) and Tik Tok (10%) as it can be seen in Figure 14. Twitter (3.6%), Pinterest (3.6%) and Snapchat (0%) were the least heavily used ones and the ones with the least users in general. This frequency analysis matches the literature review regarding the most used social media apps in Greece (Appendix subsection A.1) [18] and validates the selection of Instagram and Pinterest for the test of our third hypothesis (collective muscle memory).

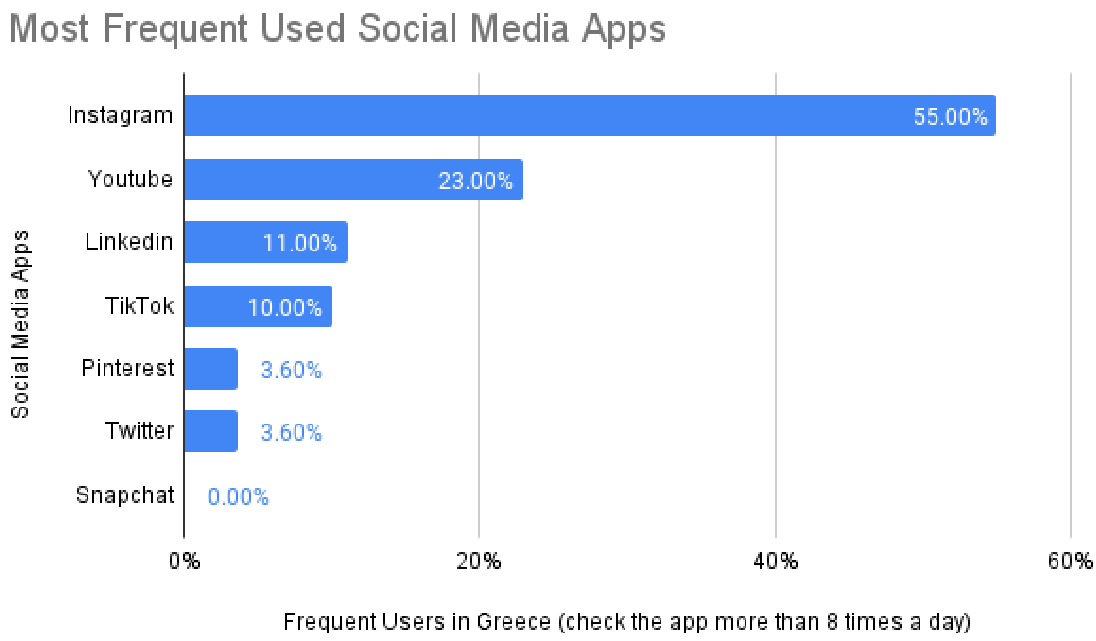

Figure 14: Most Frequent Used Social Media Apps in Greece

4.2 Hypothesis Analysis

This subsection focused on reporting the statistical tests we performed to validate our hypothesis. As we have already mentioned, there were three main hypotheses, with the first one to be broken into three simpler ones, which would also be the structure of this subsection.

4.2.1 Hypothesis 1.A.

Higher Instagram usage frequency will be translated to faster execution of the tasks on the Greek interface known to them.

For this hypothesis, a Pearson correlation coefficient was computed to assess the linear relationship between Instagram usage frequency and task performance time. On top of that, we performed a mean comparison of the time to perform the task between the Instagram usage frequency groups: rare users, less frequent users, frequent users, high frequent users and heavy users.

The Pearson correlation coefficient showed that there was a negative correlation between the two variables, $r(53) = -.29$, $p = .030$. Meaning that users with higher IG usage frequency have less execution time for the second task - click on the search button of the tab bar menu.

This negative correlation can also be observed in the mean comparison in the Table 1:

4.2.2 Hypothesis 1.B.

IG Usage Frequency	Mean Time (sec)	N	Std. Deviation
Rare Users	2.65	8	2.97
Less frequent users	1.84	15	1.14
Frequent users	1.70	13	.84
High Frequent users	1.31	12	.41
Heavy user	1.28	7	.45

Table 1: Pearson Correlation H1.A.

Frequent Instagram users will have more accidental clicks on the Dutch interface, new to them, than the less frequent users.

A one-way ANOVA was used to determine how Instagram usage frequency affects accidental clicks on the altered interface. We expected frequent Instagram users to have more accidental clicks driven by their existing muscle memory.

The ANOVA showed that there is no statistical difference in the accidental clicks on the altered interface task between the Instagram usage frequency profiles [$F(4, 50) = 1.27$, non-significant].

However, based on the recordings, more than 25% of our sample were driven by their muscle memory, and they clicked on the position where the "Create" button used to be (the new Reels button). That proved our initial assumption that users will click on a button, not because they have recognised the button (icon) but because they were expecting that clicking in that area would have the same result as all the other times they have done so, in simple words the definition of muscle memory.

On top of that, we observed on the recordings that some participants (18) used two alternative ways to create a post other than the tab bar menu. Either from the top left corner of the home screen or by entering their profile and then again in the top left corner. As a result, these participants developed different muscle memory for Task 3 (creation of a post), and the change in the altered interface has not affected them.

4.2.3 Hypothesis 1.C.

The participants who already use the altered interface (dutch population) will be faster (less task execution time) and more accurate (less accidental clicks) due to muscle memory compared with the participants who use the Greek interface.

As we stated previously, this hypothesis was used to set the baseline for the experiment. In simpler words, we needed to test that any changes in execution time and accuracy on the 3rd task (Post creation on the altered interface) would be related to the existence of muscle memory and not to bad design choices. By that, we mean that the new position of the create button (+) is still reachable with only one click; for example, it is not hidden in a menu or on a secondary screen. Also, it can be clicked in a reasonable time, meaning that the participants can easily identify it (same icon as before), or the size is still the same, so they do not have any difficulty clicking on it.

For this hypothesis, two independent T-tests for task execution time and task execution accuracy were performed between the Greek participants (with no muscle memory on the altered interface) and the Dutch participants (having muscle memory on the altered interface, as is the only existing interface in The Netherlands for more than a year now).

The first independent t-test showed that Dutch participants that were familiar with the altered interface had statistically significantly lower execution time M(2.12) sec, SD(0.22) compared to the Greek participants using for the first time the altered interface M(9.71) sec, SD(9.15), $t(53)=5.839$, $p < 0.001$.

Moreover, the second independent t-test showed that Dutch participants that were familiar with the altered interface made statistically significantly lower clicks, M(1) clicks, SD(0.00), and if we want to be precise, all of them did just one click compared to the Greek participants using for the first time the altered interface M(2.92) clicks, SD(2.42), $t(50)=5.682$, $p < 0.001$.

In conclusion, the miss clicks and longer execution time of the Greek participants were due to the lack of muscle memory and not due to bad design (not visible, needing more than one click to reach the button, weird icon selection or size) as we saw from the comparison with the Dutch sample (that was using already the altered interface for more than a year now).

4.2.4 Hypothesis 2

Social Media Apps' technique of using Muscle Memory to promote New Features is perceived by the users as a Dark Pattern.

To test this hypothesis, we conducted a frequency analysis of the responses on the following 7-point Likert scale questions. The participants were asked these questions right after completing the first three tasks. More specifically, after the 3rd task, where the participants were introduced to the altered interface of Instagram as the goal was to collect the perceived feelings.

- Instagram tricked me into performing certain actions that I did not intend to do. [Strongly Disagree - Strongly Agree]
- I felt deceived/misled by Instagram. [Strongly Disagree - Strongly Agree]

On the first question, 49% of the participants expressed the opinion that Instagram tricked them into performing specific actions they did not intend to do [Agree, Strongly Agree].

However, the percentage is rising to almost 80% for participants that had built muscle memory using the (+) button on the tab bar menu (proven clicked on the "Reels" button)

On the second question, almost 30% of the participants declared that Instagram tried to mislead them [Agree, Strongly Agree]. Most of them (31%) were neutral to that question, probably due to an unclear opinion.

In overall conclusion, participants have muscle memory using the Create button (+) to create posts. These participants, driven by that muscle memory, clicked on where they supposed the Create button (+) would be placed, but instead, they landed in the new Reels feature. So, they have first performed an action they did not want to do, and second, they felt that they had been tricked and misled. Based on these two findings, we can confirm that social media apps' technique of using Muscle Memory to promote New Features is perceived by the participants as a Dark Pattern.

4.2.5 Disguised Ads

Instagram has been criticised a lot regarding how it informs its users about the ads policy in the app, which does not give the option to opt-out of seeing ads [27]. Hence, we decided to assess participants' reactions to the sponsored content added to their news feed and see if (a) they spotted the sponsored content and (b) their opinion about such content. We did that using the questions presented in the experiment design section (point 6).

An analysis of the questions related to the disguised ads (presented earlier in 3.2), showed that almost 70% of the participants did not understand the existence of sponsored content. Moreover, 37% of the participants declared that this was very annoying. On the other hand, 41% expressed the opinion that they did not care about that. That can be explained by the fact that more than 80% of them stated that they see that kind of content frequently (28%) to very frequently (54%). They probably get used to dealing with sponsored content on their apps.

Furthermore, more than 55% of participants declared that they had not been informed or given consent to be exposed to sponsored content. Based on these findings, disguised ads are one more design pattern that fulfils the requirements to be characterised as a dark pattern.

4.2.6 Hypothesis 3

Muscle Memory acquired from one Social Media App can be applied to another Social Media app creating a phenomenon of Collective Muscle Memory.

To test this hypothesis, we filtered out all participants from the initial participants' pool that did not have Pinterest. Then the new pool (30 participants) was split into two segments, one that contains frequent, high and heavy frequent users of other social media apps (16 participants) and another one with 14 participants with less frequent users or rare ones. We assumed that the first group would perform the tasks on Pinterest (click on the search button) faster than the second one due to the existing strong muscle memory of a similar social media app (Instagram). To examine that, we performed an independent t-test on task execution time between these two groups.

The independent t-test showed that the frequent, high frequent and heavy users group that had muscle memory on the tab bar menu of Instagram has statistically significantly lower execution time M(1.31) sec, SD(0.43) compared to the less frequent and rare users group with no related muscle memory M(1.59) sec SD(0.44), $t(28)=-1.755$, $p=0.045$.

5 Conclusion

Billions of people worldwide use social media apps to make connections, communicate, create and share content online. More than half of the world now uses social media apps frequently. It is interesting to note that the average daily time spent using social media is 2h 27m. By spending that much time on social media apps, users have started building muscle memory within these apps. They can perform actions automatically, without even thinking or seeing the screen. We validated that by conducting an experiment in a mock-up of the Instagram app. We found that higher Instagram usage frequency builds stronger muscle memory. We did that through a user study where we evaluated their muscle memory using the tab bar menu on Instagram by measuring their task execution time.

Social media apps take benefit of that muscle memory by designing their interfaces in a way that minimises the cognitive load of their users. However, there have been cases where social media apps took advantage of that muscle memory, not for the convenience of the users but instead in order to promote their new features. For example, Instagram in many European countries changed the tab bar menu and replaced the probably most used "create" action item with the new feature "Reels". The new feature "Reels" contains short entertainment videos from random content creators or Instagram users. By replacing the most commonly used functions with 'new' functions at the exact location, Instagram expects the users to press the button they actually did not intend to - which is exactly what a dark pattern is by definition.

In this thesis, we simulated the release of the "Reels" feature to the Greek population. We did that because, in Greece, Instagram has not yet launched the "Reels" feature that we described previously. So we examined how that change will affect the users' experience. On top of that, we analysed how the users perceive this change. We found that the Greek users needed more time and had more miss clicks on their attempts to perform the "create" task on the interface containing the Reels feature. In combination with the previous results, their perceived feelings after these tasks lead us to conclude that the tactic of taking advantage of users' muscle memory from Instagram to promote a new feature can be characterised as a dark pattern.

Furthermore, we examined one more form of a dark pattern, the disguised ads, and we found that disguised ads are used on social media apps and are used a lot. Social media app users have been familiarised with ads in this kind of app. Finally, disguised ads are very well hidden and difficult to identify as such.
Lastly, we introduced a new term, "collective muscle memory", a phenomenon where users apply muscle memory built from one social media to another one. We examined that term by testing whether Instagram users can benefit from their muscle memory to perform similar tasks on the Pinterest app. To do so, we examined Instagram users without prior experience with the Pinterest app. We found that the frequent Instagram users with muscle memory on the tab bar menu of Instagram have statistically significantly lower execution time on Pinterest tasks than the less frequent users with no related muscle memory. Pointing out that "collective muscle memory" is a valid term in the domain of social media apps.

5.1 Limitations

As with every research, we also had some limitations and proposals for future work.

Starting with the limitations, during the experiment, we observed that some of the users were holding the device with both hands, losing part of their muscle memory (natural movement of thumb). By that, we mean that users operate their mobile phones using only one hand, and most of their actions are performed using their thumb. We believe that performing the experiment with both hands might affect their muscle memory. A proposal is to ask the users to perform the experiment holding the device with only one hand and, more specifically, the one they usually use for that action.

Moreover, remotely moderated experiments usually face difficulties with internet speed and user device characteristics. Having an effect on the execution time of the tasks and sometimes resulting in dropouts. Performing the experiment in a controlled environment where the internet would be stable and fast, and the device of the experiment would have some minimum requirements will contribute to better results.

5.2 Future work

Regarding the points for future work, there are two main directions that we are proposing. The one has to do with researching more types of dark patterns in social media apps. As we saw from the literature and our experiment on Instagram, social media apps use various types of dark patterns to increase their user base and promote their goals. It would be very interesting to explore other forms of dark patterns on Instagram or, in general, other types of dark patterns on other social media apps.

The other one concerns the new term "Collective muscle memory" that we proposed. Further research can be performed to examine how collective muscle memory can be applied in other places except for the tab bar menu. As we have seen within the social media apps industry, many common design patterns exist. Also, an interesting point would be to examine this term outside the social media apps domain and see how it can be generalised for a wider domain of apps.

Finally, in our research, we focused on our tracking mechanisms to capture the time and the taps of the participants on their mobile screens. Mobile phone eye-tracking mechanisms can provide further insights into understanding users' behaviour. However, to have that option, the whole experiment should be transferred to a laboratory where a more controlled environment will be available. That will give the ability to fulfil the main requirement for eye tracking on mobile devices, which is that it needs either the device or the head to be stable.

References

[1] Mahmoudi Sidi Ahmed et al. "Detection of abnormal motions in video". In: *Chania ICMI-MIAUCE* 8 (2008).

[2] KB Bergman. "Dark Patterns: Malicious Interface Design From A Users'perspective". MA thesis. 2021.

[3] Maartje Boer et al. "Social media use intensity, social media use problems, and mental health among adolescents: Investigating directionality and mediating processes". In: *Computers in Human Behavior* 116 (2021), p. 106645.

[4] Christoph Bösch et al. "Tales from the Dark Side: Privacy Dark Strategies and Privacy Dark Patterns." In: *Proc. Priv. Enhancing Technol.* 2016.4 (2016), pp. 237–254.

[5] Harry Brignull et al. "Dark Patterns-User Interfaces Designed to Trick People". In: *Retrieved November* 30 (2015), p. 2021.

[6] Corina Cara et al. "Dark patterns in the media: A systematic review". In: *Network Intelligence Studies* 7.14 (2019), pp. 105–113.

[7] Ming Ki Chong and Gary Marsden. "Exploring the use of discrete gestures for authentication". In: *IFIP Conference on Human-Computer Interaction.* Springer. 2009, pp. 205–213.

[8] Adobe Creative Cloud. *Adobe XD.* 2021. URL: https://www.adobe.com/creativecloud.html.

[9] Ayse Leyla Crossland et al. "The impact of interaction mechanisms with in-vehicle touch screens on task performance". In: (2019).

[10] Linda Di Geronimo et al. “UI dark patterns and where to find them: a study on mobile applications and user perception”. In: *Proceedings of the 2020 CHI conference on human factors in computing systems*. 2020, pp. 1–14.

[11] EDPB. “Guidelines 3/2022 on Dark patterns in social media platform interfaces: How to recognise and avoid them. Version 1.0”. In: (2022), p. 32022. URL: https://edpb.europa.eu/our-work-tools/documents/public-consultations/2022/guidelines-32022-dark-patterns-social-media_de.

[12] Figma. *Figma: The collaborative interface design tool.* 2022. URL: https://www.figma.com/.

[13] Daniel Franzmann, Lukas Fischer, and Roland Holten. “The influence of design updates on users: the case of snapchat”. In: *Proceedings of the 52nd Hawaii International Conference on System Sciences*. 2019.

[14] Kristian Gundersen. “Muscle memory and a new cellular model for muscle atrophy and hypertrophy”. In: *Journal of Experimental Biology* 219.2 (2016), pp. 235–242.

[15] Aurora Harley. *Don’t Prioritize Efficiency Over Expectations.* 2015. URL: https://www.nngroup.com/articles/efficiency-vs-expectations/%20www.nngroup.com/articles/efficiency-vs-expectations/ (visited on 07/13/2022).

[16] Thomas Joy. *Good User Interface Doesn’t Equal Good User Experience - Beacon Agency.* 2022. URL: https://beaconagency.co.uk/post/good-user-interface-doesnt-equal-good-user-experience (visited on 07/13/2022).

[17] Karel. *10 User Interface Design Best Practices In 2022.* 2022. URL: https://wiredelta.com/.

[18] Simon Kemp. *Digital 2022: Greece — DataReportal – Global Digital Insights.* 2022. URL: https://datareportal.com/reports/digital-2022-greece (visited on 07/21/2022).

[19] Lili Láng and Paula Dana Pudane. *Deceptive Interfaces: A case study on Amazon’s account deletion navigation and its effects on user experience.* 2019.

[20] Otto Lappi. “Eye tracking in the wild: the good, the bad and the ugly”. In: *Journal of Eye Movement Research* 8.5 (2015).

[21] Yiqin Lu et al. “Blindtype: Eyes-free text entry on handheld touchpad by leveraging thumb’s muscle memory”. In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1.2 (2017), pp. 1–24.

[22] Jamie Luguri and Lior Jacob Strahilevitz. “Shining a light on dark patterns”. In: *Journal of Legal Analysis* 13.1 (2021), pp. 43–109.

[23] Aditi M. Bhoot, Mayuri A. Shinde, and Wricha P. Mishra. "Towards the identification of dark patterns: An analysis based on end-user reactions". In: *IndiaHCI'20: Proceedings of the 11th Indian Conference on Human-Computer Interaction*. 2020, pp. 24–33.

[24] Maximilian Maier. *Dark patterns–An end user perspective*. 2019.

[25] 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*. 2021, pp. 1–18.

[26] Maze. *Maze — Product research platform for modern teams*. 2022. URL: https://maze.co/ (visited on 05/12/2022).

[27] Sara Morrison. *Dark patterns, the tricks websites use to make you say yes, explained*. 2021. URL: https://www.vox.com/recode/22351108/dark-patterns-ui-web-design-privacy.

[28] RealEye. *RealEye*. 2022. URL: https://www.realeye.io/.

[29] Neil Richards and Woodrow Hartzog. "The pathologies of digital consent". In: *Wash. UL Rev.* 96 (2018), p. 1461.

[30] Richard Shusterman. "Muscle memory and the somaesthetic pathologies of everyday life". In: *Human Movement* 12.1 (2011), pp. 4–15.

[31] Sketch. *The Best platform for digital product design*. 2022. URL: https://www.sketch.com/design/ (visited on 07/13/2022).

[32] Useberry. *Useberry: Codeless prototype analytics*. 2022. URL: https://www.useberry.com/.

[33] UXtweak. *UXtweak: Powerful tools for UX research user testing*. 2022. URL: https://www.uxtweak.com/.

[34] Rosa Van Koningsbruggen, Bart Hengeveld, and Jason Alexander. "Understanding the Design Space of Embodied Passwords based on Muscle Memory". In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 2021, pp. 1–13.

[35] Maria G Wadlow et al. "User Interface Guidelines for the Andrew System". In: *Information Technology Center* (1988).

[36] Ari Ezra Waldman. "Cognitive biases, dark patterns, and the 'privacy paradox'". In: *Current opinion in psychology* 31 (2020), pp. 105–109.

[37] Optimal Workshop. *Optimal Workshop: User Experience (UX) Research Platform*. 2022. URL: https://www.optimalworkshop.com/.

[38] Erdem Ata Yildirim. “Dark Patterns in Social Media: An analysis of Instagram’s interface [Unpublished master’s thesis]”. MA thesis. 2021.

A Appendices

A.1 Social Media Usage in Greece

Social Media App	Users in Greece (M)
Youtube	7.4
Facebook	5.15
Instagram	4.35
TikTok	2.35
Linkedin	1.9
Pinterest	1.7
Snapchat	0.86
Twitter	0.71

Table 2: Social Media Usage in Greece.

A.2 Examples of the Experiment

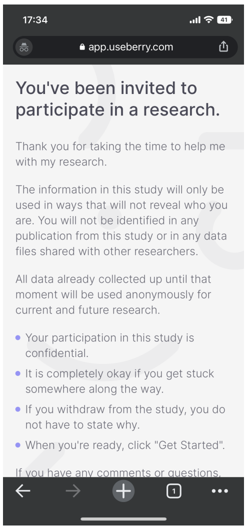

Figure 15: Welcome Screen of Experiment

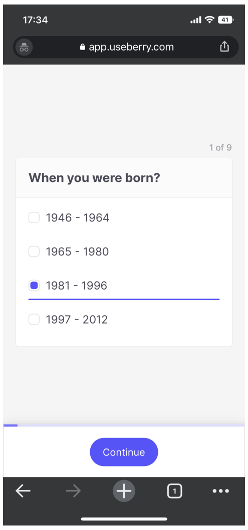

Figure 16: Demographic Question

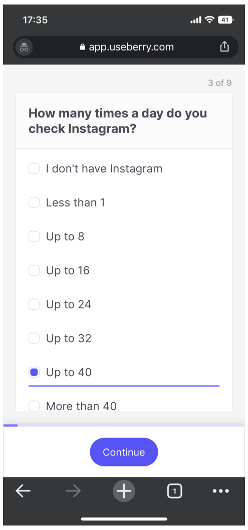

Figure 17: Social Media Usage Question

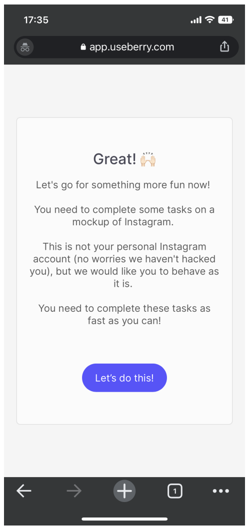

Figure 18: Guidelines for the next task

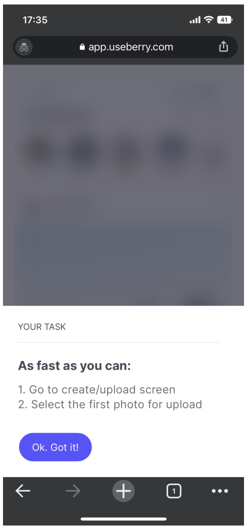

Figure 19: Example of Instagram Task

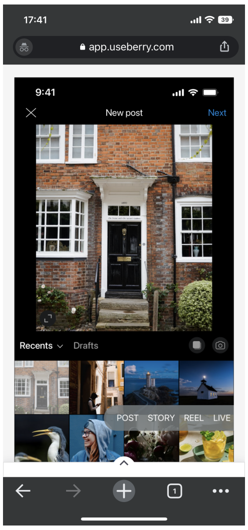

Figure 20: Example Instagram Screen

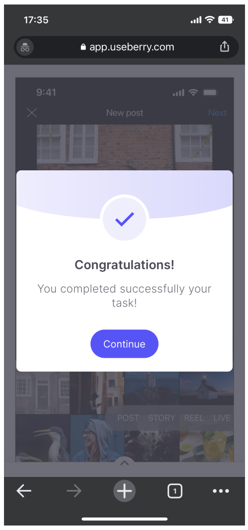

Figure 21: Complete the Task

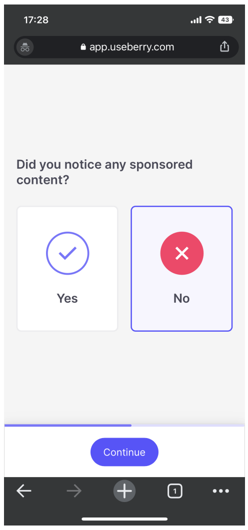

Figure 22: Question for Disguised Ads

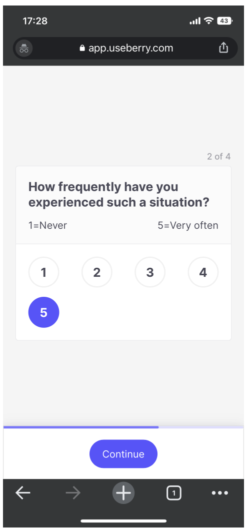

Figure 23: Question related to Disguised Ads

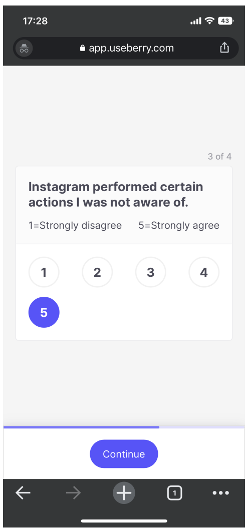

Figure 24: System Darkness Scale Question

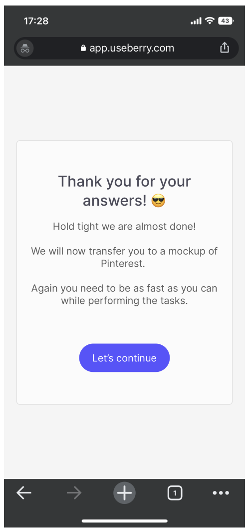

Figure 25: Introduce Pinterest UI

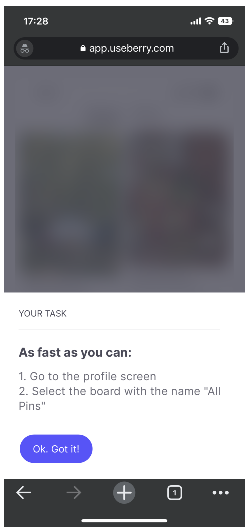

Figure 26: Example Pinterest Task

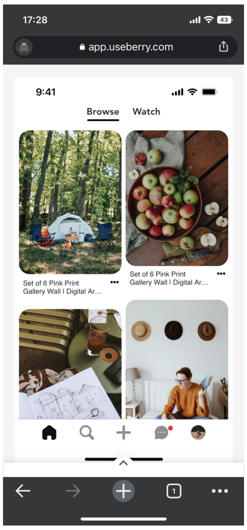

Figure 27: Example Pinterest Screen

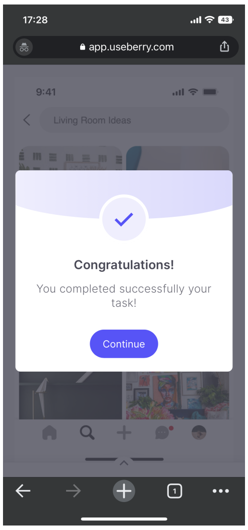

Figure 28: Complete the Task

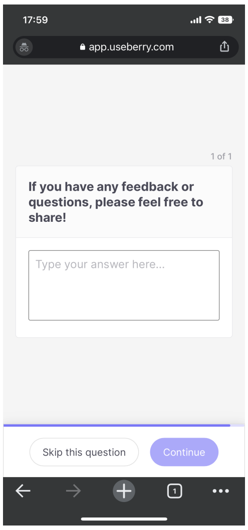

Figure 29: Ask for Feedback Screen

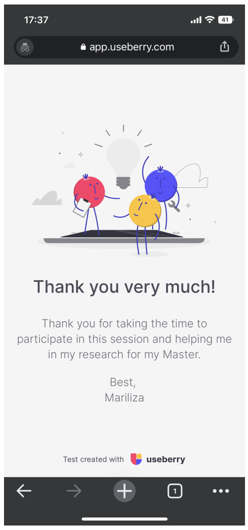

Figure 30: Last Screen of Experiment