Illuminating Muscle Memory's Sinister Side: A Social Media Case Study

Mariliza Kontogeorgou
m.e.kontogeorgou@students.uu.nl>
Utrecht University, Dept. of Information and Computing Sciences
Utrecht, the Netherlands

Christof van Nimwegen
c.vannimwegen@uu.nl
Utrecht University, Dept. of Information and Computing Sciences
Utrecht, the Netherlands

Alkim Almila Akdag Salah
a.a.akdag@uu.nl
Utrecht University, Dept. of Information and Computing Sciences
Utrecht, the Netherlands

ABSTRACT

When a task is repeated, it becomes part of procedural memory. This type of memory dedicated to movement is called "muscle memory", which allows one to perform actions unconsciously. Within the context of social media, muscle memory builds up if one uses SM applications frequently. In this paper, we investigate the effects of muscle memory within Instagram, and report the following findings: We designed a user study examining speed and accuracy of using a newly changed interface which showed slower reaction time and more errors. Combining these results with users' perceived feelings lead us to conclude that in specific UX interface changes muscle memory can be applied as a dark pattern.

CCS CONCEPTS

• **Human-centered computing → User studies**.

KEYWORDS

Deceptive Design, Social Media Applications, Instagram, Muscle Memory

ACM Reference Format:
Mariliza Kontogeorgou, Christof van Nimwegen, and Alkim Almila Akdag Salah. 2023. Illuminating Muscle Memory's Sinister Side: A Social Media Case Study. In *European Conference in Cognitive Ergonomics (ECCE '23), September 19–22, 2023, Swansea, United Kingdom.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3605655.3605664

1 INTRODUCTION

Social media applications have implemented dark patterns unique to social media ecosystems. In this paper, we document the misuse of muscle memory in social media mobile apps as a dark pattern. Muscle memory was coined as one of the cognitive traits exploited for a specific dark pattern, i.e. the bait and switch pattern [6]. Since then, it has been mentioned as a potential explanation for dark pattern use in two games [14], and one social media application [2]. However, in none of these instances, an observational study is carried out, and only the idea of how muscle memory can be exploited is sketched. This is the first research project devoted to such an observational study.

As a case study, we focus on the interface change introduced by Instagram two years ago. Through this study we show [1] how muscle memory becomes stronger when navigating the interface of a mobile app with frequent use, and [2] how a change in the UI can be misused to deceive the users to take unintended actions. Based on the results we obtain, we argue that muscle memory of users, especially on a mobile phone, can be easily used for misdirection or to coax the users to take unintended actions.

This paper thus contributes to the dark patterns literature by introducing a user study on the 'misuse of muscle memory' as a dark pattern. Mobile applications make use of both of this in various scenarios, and our paper will not only start a discussion on this important human trait and how it is used within UX, but also trigger new research to understand the phenomenon and its effects better.

2 HOW TO MEASURE THE EFFECT OF MUSCLE MEMORY 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 learned through habituation, either through explicit, intentional training or simply as the result of unintentional or unconscious learning from repeated prior experiences. Such memory is often designated as "procedural memory" or "motor memory" because it enables us to perform various motor procedures or skills intuitively. 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 [13].

Muscle memory and its effect on user experiences is researched within the UX context as well. As early as in 1988, the effect of muscle memory on the user interface was already explored. This early study reported that menu UI components take advantage of muscle memory by positioning options in the same places on different menus [17].

Recent research on muscle memory focuses on whether muscle memory can be trained to carry out tasks through UX. Such experiments define the success criteria of muscle memory training on two axes, i.e. speed (how fast users perform the tasks) and accuracy (how many correct attempts they had, measured as the error rate). For example, Chong Ming Ki and Marsden Gary examined the advantages of muscle memory as a way to memorize gestures used for authentication in mobile phones. Their main tactic was to first build muscle memory with some rehearsal sessions and then test it afterwards. Their success criterion was the number of successful

ECCE '23, September 19–22, 2023, Swansea, United Kingdom

ACM ISBN 979-8-4007-0875-6/23/09…$15.00
https://doi.org/10.1145/3605655.3605664

PIN gesture retrievals (accuracy) [4]. Their subjects were given a 24-hour rehearsal period to memorize 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 practice their passwords.

In a similar context, Van Koningsbruggen et al. examined the usage of muscle memory in recalling passwords; this time using embodied passwords. Embodied interaction focuses on the creation, manipulation, and sharing of meaning through our interactions with objects [10]. The participants were asked to memorize specific movements on specific items as 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 practicing, after which they were tested twice: one day after the training and one week afterwards [15]. All participants could recall and reproduce their embodied passwords during these tests, but they were not as successful in remembering their alphanumeric passwords.

We conclude this section with the observation 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). 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), hence there will be no need for a training period.

3 SOCIAL MEDIA APPS AND MUSCLE MEMORY USE

Certain dark patterns are only found in social media platforms, one of them is the so-called "Fake Notifications". These apps send notifications without an important reason, such as "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 regularly [3]. 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 recognize them [11]. A third social media specific UX design is the so called 'infinite scroll', and as it keeps the users attention engaged longer than intended, it is cited as a dark pattern [12]. We believe that in these three cases, the muscle memory plays an important role.

We believe 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 [8]. Snapchat faced heavy criticism and complaints when they made a significant change to their interface [5]. Instagram is another example: in many European countries, Instagram's probably most used "create" action item is replaced with the new feature "Reels" [7]. The new feature "Reels" contains short entertainment videos from random content creators or Instagram users. We will focus on this replacement, and design an experiment to observe the effects of this in user behavior and perception. The rest of the paper is thus devoted to explaining our experiment setup, report our results, and conclusions.

4 RESEARCH METHODS

Literature shows that having many repetitions of the same action results in building muscle memory. In this section, we document our experiment design to test if this theory applies to the social media apps domain.

4.1 Experiment Design

As we previously mentioned, two years ago Instagram changed its tab bar menu in many European countries [7]. In Greece, Instagram had not yet launched this UX change. Thus, focusing on Greek population, we examine how the new UX affects the users behavior and perception. Figure 1 shows the Instagram interface for Greeks and Dutch population. The latter has the updated UX interface.

With the "create" action item, the users have the ability to share an image or a short video through the creation of a post or a story. The "create" action button was located in the tab bar menu's prominent position, the center. Instead of adding a sixth action item for the Reels, Instagram replaced the "create" action item with the "Reels" and moved the "create" to the top right corner. In Figure 1, the two UIs, with and without the "Reels" feature, are depicted.

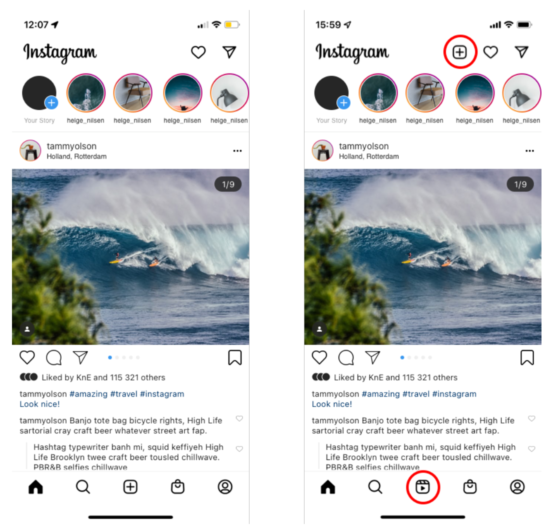

Figure 1: Left: Greek Interface - Right: Dutch Interface

Our experiment design consists of four stages: (1) Questions regarding demographic data, (2) questions to quantify participants social media usage, (3) performing tasks on the generated Instagram prototype, (4) open question to collect any additional feedback. We accepted participants with all types of social media usage experience, even ones that have never used social media apps before.

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. [1] in their research: How many times a **day** do you check the following social media apps? We selected the most used social media apps [9] that follow the same design pattern on their tab bar menu. 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 familiarization **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 familiarization 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.

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 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, i.e. 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.

We asked the participants 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. 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, by looking at how many clicks they need 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 are chosen from the System Darkness Scale proposed by [16] and adjusted to fit our case:

(1) Instagram tricked me into performing certain actions that I did not intend to do. [Strongly Disagree - Strongly Agree]

(2) I felt deceived/misled by Instagram. [Strongly Disagree - Strongly Agree]

Experiment Interface Designs and Measurements. We used the software *Sketch* to design our mock-ups and connected it to the tracking tool Useberry. We designed over 25 screens for three fully interactive prototypes of Greek & Dutch Instagram [1] From Useberry, we collected the screen captures, the task execution time (in milliseconds), the number of miss clicks, and the heatmaps (for the quantitative exploration of testers' clicks/taps). We had two questionnaires to assess the usage frequency of social media apps and the dark pattern awareness of the users. In the result section, we explain in detail the quantitative and qualitative analyses we performed on these data.

[1] 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.

5 RESULTS

We targeted the Greek population, where Instagram still has an earlier version than the one in other European countries. To create a baseline, we also had a small set of participants from the Netherlands who have been using the updated interface since a year. For these participants, we did not collect demographic data, nor their social media usage habits. We also did not submit them to questions about their perceived feelings.

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 screen recordings, these participants were facing either technical issues or were not focusing on the experiment. Our Greek sample population is 55 participants who were mainly aged between 26 and 41 years old. 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. We identified the following five profiles of usage based on the times they check a social media app per day: (1) Rare Users - Less than 1, (2) Less Frequent Users - Less than 8, (3) Frequent Users - Less than 16, (4) High Frequent Users - Less than 24, (5) Heavy Users - More than 24.

5.1 Hypothesis Analysis

Hypothesis 1: 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, $r(53) = -.29$, $p = .030$. Users with higher Instagram usage frequency have less execution time for the second task - click on the search button of the tab bar menu.

Hypothesis 2: 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 recognized 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, using their muscle memory to carry out the action.

On top of that, we observed on the screen 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.

Hypothesis 3: The participants who already use the altered interface (Dutch population) will be faster (shorter task execution time) and more accurate (fewer accidental clicks) due to their re-trained 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. Since the new interface has the following design features: (a) the new position of the create button (+) is still reachable with only one click, (b) it is not hidden in a menu or on a secondary screen, (c) it can be clicked in a reasonable time, (d) the participants can easily identify it since it has the same icon as before, (e) the size of the icon is still the same, we assumed the interface design has not a bad design choice, and hence, theoretically the participants should not have any difficulty finding and clicking on the create icon.

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 who were familiar with the altered interface neded statistically significantly shorter 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 made statistically significantly fewer clicks, M(1) clicks, SD(0.00). To be precise, all of them needed just one click compared to the Greek participants, M(2.92) clicks, SD(2.42), $t(50)=5.682$, $p < 0.001$.

6 CONCLUSION

Users who spend regular time on social media apps, build muscle memory, which in turn causes actions to be performed automatically. In this paper, we tested if such a muscle memory is used for malicious purposes. We focused on Instagram's introduction of a new feature, i.e. the "Reels" feature which replaced the most used activity, i.e. the "Create" feature. Two different sets of participant contributed to this study: Dutch and Greeks. As the changes in the Instagram interface were not introduced in Greece yet, we were able to test the differences in reaction time and accuracy between these two populations. 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. Moreover, Greeks reported feeling 'deceived' about the change in the interface.

During the analysis of the data, we observed that some participants have developed a different muscle memory than expected, which affected our results. A follow-up study should use more participants where similarly built muscle memory can be tested more adequately. Also, eye-tracking and heat maps should be used to achieve more accurate results, and checked if these approaches will bring new insights. Beside expanding experimental designs for muscle memory, future research should also focus on developing ways to examine other muscle-memory misuse cases in social media apps, such as infinite scrolling, disguised adds, and fake notifications among others.

REFERENCES

[1] Maartje Boer, Gonneke WJM Stevens, Catrin Finkenauer, Margaretha E de Looze, and Regina JJM van den Eijnden. 2021. Social media use intensity, social media use problems, and mental health among adolescents: Investigating directionality and mediating processes. *Computers in Human Behavior* 116 (2021), 106645.

[2] Wilhelmina Maria Botes, Rachele Carli, Arianna Rossi, Lorena Sanchez Chamorro, Cristiana Santos, and Anastasia Sergeeva. 2022. Feedback to the Guidelines 3/2022 on "Dark patterns in social media platform interfaces: How to recognise and avoid them". (2022).

[3] Corina Cara et al. 2019. Dark patterns in the media: A systematic review. *Network Intelligence Studies* 7, 14 (2019), 105–113.

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

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

[6] Colin M. Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L. Toombs. 2018. The dark (patterns) side of UX design. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–14.

[7] Thomas Joy. 2022. Good User Interface Doesn't Equal Good User Experience - Beacon Agency. https://beaconagency.co.uk/post/good-user-interface-doesnt-equal-good-user-experience

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

[9] Simon Kemp. 2022. Digital 2022: Greece — DataReportal – Global Digital Insights. https://datareportal.com/reports/digital-2022-greece

[10] Paul Marshall, Alissa Antle, Elise Van Den Hoven, and Yvonne Rogers. 2013. Introduction to the special issue on the theory and practice of embodied interaction in HCI and interaction design. , 3 pages.

[11] Arunesh Mathur, Mihir Kshirsagar, and Jonathan Mayer. 2021. 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*. 1–18.

[12] Thomas Mildner and Gian-Luca Savino. 2021. Ethical user interfaces: Exploring the effects of dark patterns on facebook. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–7.

[13] Richard Shusterman. 2011. Muscle memory and the somaesthetic pathologies of everyday life. *Human Movement* 12, 1 (2011), 4–15.

[14] J. H. Tiemessen. 2022. The Time is Ticking: The Effect of Deceptive Countdown Timers on Consumers' Buying Behavior and Experience. (2022).

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

[16] Christof van Nimwegen, Kristi Bergman, and Almila Akdag. 2022. Shedding light on assessing Dark Patterns: Introducing the System Darkness Scale (SDS). In *35th International BCS Human-Computer Interaction Conference 35*. 1–10.

[17] Maria G Wadlow, Christina Haas, Dan Boyarski, and Paul G Crumley. 1988. User Interface Guidelines for the Andrew System. *Information Technology Center* (1988).