Trigger-Action Programming in Digital Self-Control Tools to Improve their Long-term Effectiveness

BY

MATTIA CHIARLE
B.S., Politecnico di Torino, Turin, Italy, 2022

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Master of Science in Computer Engineering
in the Graduate College of the
University of Illinois at Chicago, 2024

Chicago, Illinois

Defense Committee:

Debaleena Chattopadhyay, Chair and Advisor

Ugo Buy

Alberto Monge Roffarello, Politecnico di Torino

ACKNOWLEDGMENTS

First, I would like to thank my advisors, Debaleena Chattopadhyay and Alberto Monge Roffarello, for their invaluable support during these months. Without all your precious suggestions and ideas, nothing of this would have been possible.

Thank you, Mom, for everything you have done for us. Among all the lessons you taught me, the most important is that being nice and doing the right thing is always the best choice, no matter the outcome. I would not be the person I am today without you.

Thank you, Dad, for showing me that hard work always pays off and teaching me to be ambitious and never settle.

Thank you, Leo and Tommi, for being my rock over the past years. Even if we live in different parts of the world, you are always in my mind, and I could not be prouder of what you are building.

Thank you, Franci, for being the best companion and support I could have dreamt of for most of this journey. You profoundly changed me in many aspects, and without you, I would not be here today. Thank you, Paola, Gino, Filippo, Diddi, Ale, Leo, Enrico, Rebecca, and all the uncles and cousins. You have been my second family.

Thank you, Ema and Leo, my lifelong friends. Even if years pass and we have different career choices, we will always be there for each other.

ACKNOWLEDGMENTS (continued)

Thank you, Nonna Lella, for being such a fantastic grandmother. I always look forward to our daily call, and my heart is full of beautiful memories with you. Thank you, Nonno Franco, for all the moments we spent together. You are always in our hearts.

Thank you, Nonna Franca, for all the beautiful days in Sangano. I always look back at them with joy.

Thank you, Gigi, for all the good you have done for Mom and us. You treated us as a part of your family from the first day, and we have never been happier. Thank you, Cucu, Morbi, Sam, and Iaia, for being part of this fantastic enlarged family.

Thank you, Chiara, for all the moments you helped me, like my high school semester in Boston.

Thank you, Carlo and Piera, for being part of our family all these years. You have always been there for us; I will never forget it.

Thank you to all the people I met at Politecnico di Torino. A special thanks to Simo and Cava, the first people I got to know after COVID, and the fantastic group of friends formed during the first year of the Master's degree. Thank you, HKN, for making me grow enormously and for all the passionate and incredible people I had the opportunity to meet.

Thank you to all the people I met in Chicago. It would not have been the same without you. Thank you, Fra, Fede, and Peppe, for being my roommates. If I could go back to one year ago, I would choose you again. Thank you to Nico, Lore, Falco, Marti, and Ali for all our funny moments, like the movie nights in Palazzina. I hope we will keep in touch after this year.

ACKNOWLEDGMENTS (continued)

Last, thank you to everyone who participated in my user study. I highly appreciate your availability and contributions, without whom this thesis would not exist today.

MC

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

CHAPTER **PAGE**

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

DSCT	Digital Self-control Tool
TAP	Trigger-Action Programming
HCI	Human-Computer Interaction
EUD	End-User Development
HAM	Habit Alteration Model
DBCI	Digital Behavior Change Interventions
UIC	University of Illinois at Chicago

SUMMARY

Smartphones have become fundamental in our everyday lives. Thanks to the enormous technological advancements of the last twenty years, we can perform a variety of tasks that would be much more complex, if not unfeasible, without them. Unfortunately, this led people to use those devices much more than needed, losing control over them. To contrast this phenomenon, researchers introduced the notion of digital wellbeing and tried to improve it with Digital Self-Control Tools. Those tools guide the user toward more conscious and healthier behaviors, but if not properly calibrated, they may cause reactance, leading to their abandonment.

This document will analyze previous research on this field and End-User Development. We will describe a user study we conducted to extract people's needs in the digital wellbeing field and their opinions about Digital Self-Control Tools and Trigger-Action Programming. We will then map our findings with a theoretical framework, the Habit Alteration Model, which will allow us to understand the study's outcomes better. Inspired by the framework and the study, we will propose an Android application based on rule creation focused on improving its long-term effectiveness. To conclude, we will explain the limitations of our work and how to improve it in the future.

CHAPTER 1

INTRODUCTION

In the last few years, smartphones have become pervasive in our everyday lives. We can do everything with them, from keeping in touch with our friends and families to using them as entertainment. Thanks to their evolution, many tasks that were difficult or even impossible not much ago became straightforward. However, its flip side is that we constantly look at these devices, even when we do not need to. As a result, many people use them much more than they would like. This excessive and uncontrolled usage causes many alarming consequences, such as psychological diseases.

Given the relevance of these phenomena, the concept of digital wellbeing was created. This new part of each individual's wellbeing targets our usage of the technology and our relationship with it. It can be achieved either autonomously, relying on our willpower, or through the assistance of external tools, named Digital Self-control Tools (DSCTs). Each DSCT implements some interventions, which, in different ways, should help the user achieve their desired type or amount of usage.

The constantly growing scale of the problem rapidly increased DSCTs' popularity so that even the most important mobile operative systems (Android and iOS) include some essential interventions nowadays. There are also many successful and widespread commercial applications, such as Forest [2].

Unfortunately, many studies revealed that, although effective in the short term, users abandon most interventions in the long term. An increasingly more popular branch of HCI started studying the reasons behind this loss of effectiveness and proposed solutions that could fix it. This thesis tries to draw on them, exploring if the End User Development paradigm can increase the long-term effectiveness of DSCTs.

1.1 Thesis Goal

This thesis investigates why users abandon or ignore their DSCTs and what measures we could take to prevent this behavior. Inspired by the literature review phase, the initial idea was that one key factor behind this trend is the reactance those tools cause in the user. Initially, thanks to curiosity and excitement for the novelty, it is not a primary concern. Still, it becomes increasingly pressing as time passes until they can no longer bore it.

To reduce the friction, we thought allowing the user to customize the behavior of these interventions fully would have allowed them to build something that fits them better. End-User Development could enable that. We initially focused on Trigger-Action Programming because different studies identified it as the most intuitive EUD technique, and there was an immediate mapping between rules and the most common interventions available.

We then tried to understand whether our hypotheses were valid by running focus groups and co-design sessions with end users. Based on their outcomes and the theoretical framework (the Habit Alteration Model), we developed a TAP-based Android application that allows users to create interventions.

1.2 Thesis structure

Chapter 2 provides a quick overview of digital wellbeing, why it has become so relevant in recent years, and some existing solutions to improve it.

Chapter 3 will describe some of the most relevant HCI research in the digital wellbeing and EUD (particularly, TAP) field.

Chapter 4 outlines a user study we conducted at the beginning of our research, concluded by an analysis of the results.

Chapter 5 reports the theoretical framework (the HAM) we used to map our findings to a behavioral theory.

Chapter 6 describes an Android application prototype we designed and developed, which should improve the long-term effectiveness of its interventions.

In Chapter 7, we will discuss the limitations of our work, suggesting possible ways of improving and continuing it.

Last, in Chapter 8, we will summarize the whole research and discuss its contribution to the digital wellbeing field.

CHAPTER 2

BACKGROUND

Digital wellbeing is a fascinating topic. Limiting our smartphone usage by installing additional tools may initially seem absurd, but nowadays, it is needed more than ever.

Google was probably the first big company to grasp the criticality of this problem. Interestingly, the understanding of how those devices, though useful, could affect people's wellbeing has its roots back to 2012 [3], when Tristan Harris, one of Google's employees, had a significant intuition: Smartphones, and in general, many applications developed for them, were not designed with users' wellbeing in mind. Unfortunately, as we can see, he did it right, and this issue worsened enormously compared to twelve years ago. Moved by his idea, he created a presentation, "Call to Minimize Distraction & Respect Users' Attention," which became quite popular within the company and eventually reached Larry Page, the CEO. His efforts made him become Google's first-ever "design ethicist."

The first big digital wellbeing breakthrough was in 2018 when Google introduced the Digital Wellbeing suite on Android. Again, its motivation can be found in another personal story, shared directly by Sameer Samat, Google's VP for Android and Google Play, during the Google I/O in which the company first presented those tools [3]. On a recent vacation, his partner, with an excuse, took his smartphone and locked it inside the hotel safe. He immediately felt angry when she told him he would get it back in seven days. However, after a few hours, he was surprisingly relieved and enjoyed the trip more than he would have with his smartphone.

One month later, Apple introduced tools designed to improve people's digital wellbeing in iOS 12 [4].

If even some of the most influential figures in companies that make profits based on how long we use their services think this is a problem too big to be ignored, the relevance and criticality of digital wellbeing are evident.

In the last few years, two contrasting trends have occurred. On one side, Digital Self-Control Tools (DSCTs) have become more relevant. They are tools installed in browsers or smartphones that include a set of interventions designed to improve the quality of the use or reduce the time spent on distracting apps or websites. There are four main categories of DSCTs, according to [5].

1. **Block/Removal**. They remove or block some features of the application.
2. **Self-Tracking**. They provide the user with information to track their usage.
3. **Goal Advancement**. They allow the user to set goals related to digital wellbeing and see if and how they are accomplishing them.
4. **Reward/Punishment**. They praise users when they behave well and punish them when they do not.

We can further divide the interventions into two main categories: awareness and blocking. The former group includes all the interventions that do not force the user in any way but provide contextual information so they can ideally make decisions they will not regret later on. One of the adopted techniques is nudges, defined as "any aspect of the choice architecture that

alters people's behavior predictably without forbidding any options or significantly changing their economic incentives" [6]. Based on that, Monge Roffarello et al. [7] developed a DSCT. The latter is probably the most common and includes all the interventions that force a user's behavior. The most famous example is usage timers, which block the application when we use it for a certain amount of time.

Today, users can find many applications to enforce interventions to limit their usage. Among the most popular commercial products, we can mention Forest (10M+ downloads on Android) [2], an application that motivates the user not to use the smartphone for a certain amount of time. It is one of the most popular digital wellbeing applications, and it was mentioned multiple times during our interviews. Another known DSCT is StayFree (10M+ downloads on Android) [8], which offers a set of interventions like usage goals, global comparisons, and time limits. One powerful option, still popular but not as known among users, is AppBlock (5M+ downloads on Android) [9]. One of its main features is the ability to set usage timers that can be locked, preventing modification during the day. Furthermore, other smartphone producers started including interventions in their devices, such as OnePlus' Zen Mode [10]. This feature completely blocks the smartphone (except for emergency calls and the camera) for a fixed amount of time the user chooses.

An increasing number of companies have begun studying and including mechanisms in their products to maximize the time users spend on their applications. The most commonly adopted strategy is implementing Dark Patterns in the interface. This term, coined by Brignull in 2010 on [11] (formerly, `darkpatterns.org`), defines them as "tricks used in websites and apps that

make you do things that you did not mean to." Considering digital wellbeing, this term refers to all the design choices made to make the users spend more time than they would want to on applications or websites. TikTok and Instagram Reels are famous examples, which adopt dark patterns like autoplay content and endless scrolling. Some studies, like [12], analyzed the dark patterns the most famous social networks exploit. Furthermore, another study pointed out that even if users are aware of them, this is not correlated to their resistance [13].

It is clear that while DSCTs can support users in their daily smartphone use, lousy design decisions partially encourage such negative usage. People would need those tools less if companies designed applications with user wellness in mind. Some encouraging signals, like Google's and Apple's efforts, are still insufficient. Until there is a shift in this trend, providing users with tools to mitigate the dire consequences of those choices is the best we can do.

CHAPTER 3

PREVIOUS WORK

At the beginning of our study, we performed a literature review to understand the state of the art and how we could contribute to it. We investigated two main topics: digital wellbeing and End User Development. As discussed in 3.3, many studies reported TAP as the most intuitive and effective EUD strategy. Thus, we focused our efforts on it for the latter area.

3.1 Digital wellbeing interventions

Digital wellbeing has become increasingly relevant in the HCI field in recent years. This trend is due to two emerging phenomena. The first is the increasing number of people who realize their excessive and unhealthy smartphone use. Studies also proved that this issue is strictly related to health problems in adolescents and young adults [14] and reduced work engagement due to sleep interference [15].

As shown by Figure 1, this topic gained considerable popularity since May 2021. Among the factors, we can link this growth to the COVID pandemic, which also emerged during our interviews as a breakthrough in users' lives and highly changed their relationship with technology. For example, P29, one of the participants of our study, noticed, "COVID and the pandemic have radically changed my use of digital devices, social media, and daily schedules. I have never been able to go back to how it was before, and I have worsened compared to before. I realized I use these devices more. For example, when I went to study rooms before

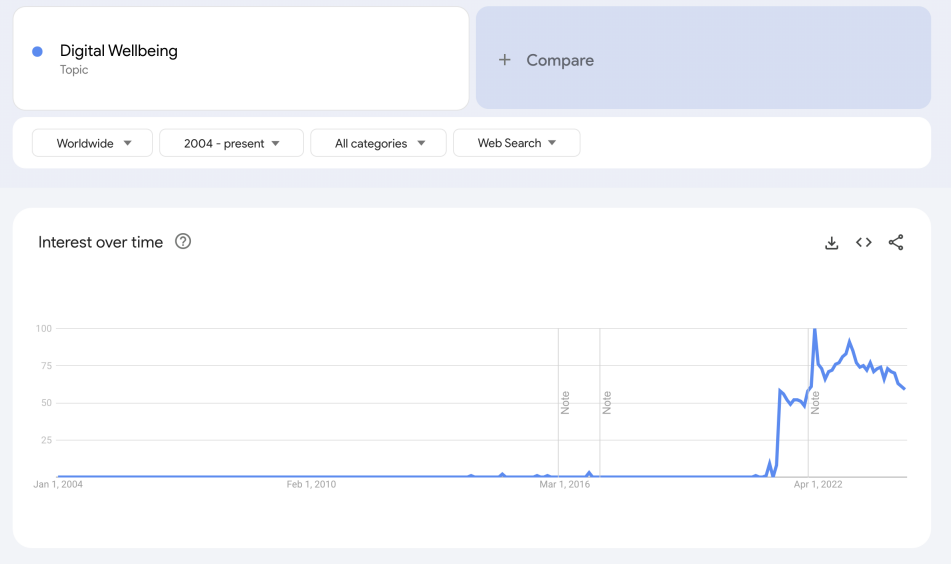


Figure 1: The popularity of the search term 'Digital wellbeing' from 2004 to today, according to [1].

the pandemic, I only used the smartphone during lunch breaks, while today, I use it much more (every hour when I take a break)."

The second one is the evidence that many people abandoned their DSCTs after some time. While HCI studies have limited potential to improve the popularity of digital wellbeing and DSCTs, they can analyze what does not work in them and how to improve their long-term effectiveness. They can then define design guidelines, which DSCTs creators can use to enhance their products. Also, they can seek unsatisfied user needs or existing issues in this field and propose new interventions.

Throughout the years, researchers have tried different paths to reach those results. Ulrik Lyngs et al. [16] analyzed the Problematic Facebook Use (PFU), describing the existing prob-

lems and proposing two interventions - one awareness-related, one block-related - to mitigate them. Although they focused on Facebook, we can extend some ideas to other similar social media. One interesting aspect was the distinction between active and passive use. They described active use as "creating content and engaging with friends," while passive as "consuming content created by others without actively engaging." The most concerning type is the last one, as different studies found a correlation between it and a negative user's wellbeing [17; 18]. The awareness intervention consisted of prompting the user for their goal when visiting Facebook and periodically reminding them of it. The blocking one, instead, removed the newsfeed. Based on the results of their experiment and the theory behind digital behavior change interventions, they expect the blocking intervention to be more effective in the long term. However, in their study, both the alternatives positively affected perceived control and behavior. Furthermore, they found that people differ in their Facebook usage, influencing what they seek as interventions and how they perceive them. They also noticed that if people are not entirely in control of the intervention, they may be annoyed and ignore or abandon it.

Holte et al. [19] argue, in their research, that it is essential for the interventions not to be restrictive, i.e., they should not be blocking but give the user full autonomy. They then tested the effect of turning the smartphone display grayscale. In this way, the smartphone should be less appealing, thus reducing the negative usage. In the study, participants with this condition tended to be less compelled to check their smartphones and enjoyed it less. They also had an anxiety reduction, although there was no significant difference concerning the number of phone checks.

Davis et al. [20] conducted a study on one of the populations most in danger when discussing digital wellbeing - teenagers. Prior studies have shown that many teenagers lack control over social media usage, which can negatively impact their well-being. As a result, they developed Locus to address this issue. Locus is an application that prompts them to reflect on their intentions each time they start a social media usage session. Furthermore, it delivers a notification at the end of the day, making users reflect on their daily usage and asking about their goals for the next day. To design it, they ran co-design sessions with some teenagers, during which contrasting needs emerged, primarily related to the prompt frequency. Those disagreements also appeared during the open trial field deployment, which they used to evaluate the effectiveness of the application. The different needs led to a request for customization options, which were essential for the application. Overall, Locus helped the participants think about their social media usage goals. It also increased their self-control and autonomy and decreased their absentmindedness.

Last, Ko et al. [21] had a motivation similar to our study. They realized users often have difficulties maintaining their strategies to limit smartphone usage. To improve this aspect, they designed NUGU, an application based on the social cognitive theory. NUGU is a group-based intervention composed of three parts. The first, self-monitoring, gives users an overview of their usage for a specific day. The second is goal-setting, which works similarly to OnePlus' Zen Mode. Last, the social learning and competition aspects included in the application allow the users to compare themselves with their friends. In this way, they should be motivated to use their smartphones less. Thanks to a comparison with an alternative version of NUGU,

which missed the last feature, they observed that the social aspect motivated users to set goals in more contexts and with longer goal time. The effectiveness depended also on how close the group members were: the closer, the better the result. Again, users also expressed a need for customization. In particular, they wanted to allow the usage of some applications, such as the notes or the dictionary app, during their missions.

3.2 DSCTs design

Until now, we have analyzed studies aimed at creating new interventions. Other researchers, instead, analyzed the available interventions, focusing on their weaknesses to share design guidelines or considerations to improve DSCTs.

The most complete work is probably the systematic review and meta-analysis conducted by Monge Roffarello et al. [5]. In it, they examined 62 papers (filtered out from a starting base of more than 4800) related to digital wellbeing. They classified them based on different factors, such as motivations and research goals, the interventions category (block/removal, self-tracking, goal-advancement, and reward/punishment), or ethical factors. For us, the most relevant part is the evaluation of the existing DSCTs. They reported that, analyzing all the studies that included a withdrawal phase (i.e., a phase in which the users remove the interventions), the behaviors formed when the tool is active tend to disappear when they do not use it anymore. This trend can be explained by DSCTs primarily being based on self-monitoring.

To improve them, the authors propose some guidelines. First, we should stop focusing on technology overuse. Digital wellbeing is much more than that; it includes many fundamental aspects, such as contrasting dark patterns or reducing distractions. Those tools should be

helpful for everyone and not just for the users who think they are using their smartphones too much. Also, there should be mechanisms to ease the self-programmed part of the intervention. DSCTs should provide aid and guidance to users, suggesting the ideal configuration for them and adapting as their use changes. They also indicate that, although challenging and often unfeasible, redesigning the interface of applications could support users. They conclude by reporting strategies to assess the effectiveness of the designed interventions, like basing design decisions on behavioral theories and conducting more prolonged user studies to test their efficacy in the long term.

Schwartz et al. [22] proposed four design patterns to reduce the abandonment of DSCTs.

- **Continuously Variable Interventions**. Instead of having only the on and off modes, the interventions should scale across multiple intermediate levels. This idea would help minimize the user friction caused by their effects.
- **Anti-Aging Design**. DSCTs should not be vulnerable to software aging, i.e., they should keep working no matter the changes in the underlying application or platform.
- **Obligatory Bundling of Interventions**. Instead of allowing the user to use a single intervention, DSCTs should force them to adopt a set of interventions. This solution would allow us to reduce the risk of abandonment due to the failure of a single intervention.
- **Intermediary Control Systems**. An intermediate control system should suggest or automatically manage the active interventions for users. A system like this would avoid disappointment due to a wrong or ineffective configuration.

They also developed a DSCT, Time Sidekick, which implemented those principles.

To conclude, Lukoff et al. [23] used the Self-Determination Theory to explain why users are hostile toward interventions they decided to enable when enforced. The problem with many DSCTs is that they undermine the user's autonomy, i.e., the user feels they must use the tool rather than want to. Those tools should adapt to the current situation to maximize autonomy, and, more importantly, they should adopt customization rather than personalization. They note this contradicts the common UX design principles that try to reduce user reflection as much as possible. Still, given the peculiar nature of DSCTs, an explicit commitment could reduce the friction.

Furthermore, the base problem, as described by Figure 2, is that, when creating an intervention, the user delegates their goals to the tool. However, the time of this decision and the time of the intervention enforcement are different, and the user may have forgotten why they did it. Thus, reminders of the original motivation when the DSCT enforces the goal or similar could be beneficial.

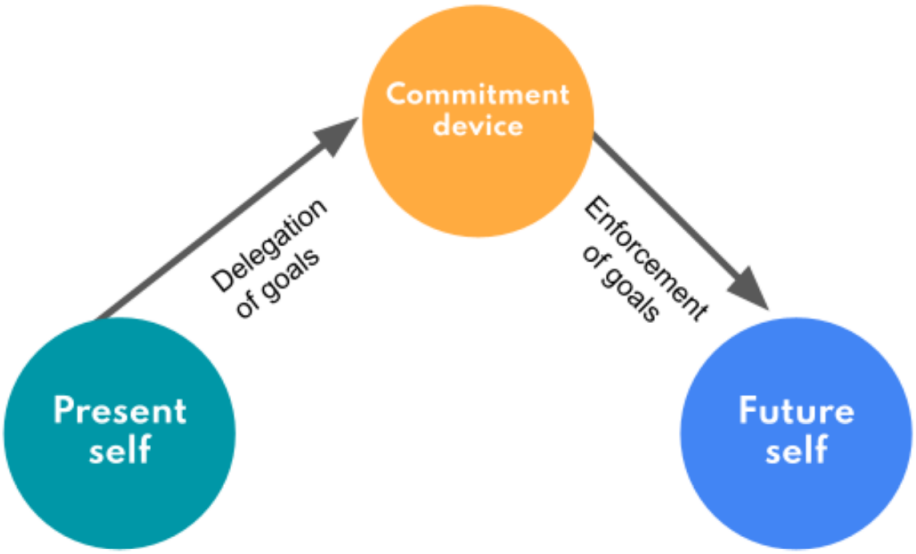


Figure 2: The possible cause of friction, according to Lukoff et al., in intervention enforcement.

3.3 End User Development

End User Development is an exciting paradigm. Lieberman et al. [24] defined it as "a set of methods, techniques, and tools that allow users of software systems, who are acting as non-professional software developers, at some point to create, modify, or extend a software artifact." As explored by [25], there are different ways to implement EUD. However, one paradigm that has acquired increasing popularity, especially in the IoT field, is Trigger-Action Programming.

Given the novelty of this approach, researchers conducted multiple studies to understand how users perceived rule creation and how to simplify this process for them. One critical aspect was bugs. Since TAP, although relatively primitive, is still a form of programming, users can

introduce bugs. For this reason, some studies analyzed if debugging is needed and how to implement it [26; 27]. Others focused on whether users can identify bugs on their own and fix them [28].

Corno et al. proposed an ontological framework that automatically adapts rules to the context, thus reducing the number of rules a user must create to enforce a specific behavior [29]. The idea is that the user expresses rules in the form of what they want to achieve instead of individually creating rules for each smart appliance they own. Thanks to an IoT ecosystem model and context, the framework will automatically determine how to enforce the desired behavior. Based on it, they developed a dynamic recommendation system to support users in the rule creation process [30] and HeyTAP. This conversational TAP platform can translate users' abstract needs into actual rules [31].

Huang et al. [32] further classified triggers and actions into additional categories. Triggers can be events (instantaneous events) or states (a boolean condition that can be checked at any time). Actions, instead, can be instantaneous (an action composed only of one step), extended (actions that take some time to complete, but when they end, they revert to the original state), or sustained (actions that will not revert automatically). They then analyzed which categories IFTTT, a popular TAP-based application, offers. They noticed it includes only event triggers (the state triggers are mapped into events). Concerning actions, instead, all three categories are supported without distinguishing between them.

To understand if and how this affected users, they ran two studies, one related to the interpretation of existing rules and one related to their creation. They found inaccuracies in

the participants' mental model and proposed improvements to TAP interfaces to mitigate this issue.

Zhao et al. [33] designed and evaluated new TAP interfaces, based on the analysis of the existing ones, to compare rule-based programs semantically and not just syntactically. In particular, they measured a control condition (rules, a standard rule creation interface with no additional aids, and text-diff, highlighting the differences between rules text) with two novelties. The first, *Outcome-Diff*, compares two rules by showing the situations in which the rules would provide different outputs. It has two variants: *Flowcharts* (it shows all the situations in which the rules provide different outputs) and *Questions* (it asks the user which outcomes are desirable in specific situations). The second one, *Property-Diff*, compares high-level rule properties. In this way, the user better understands higher-level behaviors in their systems. After a user study, they discovered that, for simple programs, all the interfaces are effective. However, their novel interfaces outperformed the control condition as the rules became more complicated.

Each new interface was more effective in a specific context. For example, Property-Diff was the best when the participants had to reason about abstract differences.

To conclude, Ur et al. [34] extracted and analyzed more than 200,000 IFTTT programs (also called recipes). One of the most compelling findings was that most authors create only a few recipes. Furthermore, many recipes were duplicated, suggesting that users prefer to make the program themselves rather than searching for something already existing. For this reason, it is unsurprising that more than half of the recipes are used by only one user. Those insights

show that while a collection of existing programs is helpful for some, most prefer creating the programs they need.

CHAPTER 4

USER STUDY

This chapter describes the whole study, from the design phase to the analysis of the gathered data.

4.1 Goal

As we have seen, the thesis aims to understand why DSCTs tend to be abandoned in the long term and propose a solution that addresses this problem. Thus, inspired by the works described in Chapter 3, we decided to run a qualitative user study. Our goal, which we will better explain in the following sections, was to directly understand from the users why they quit using their DSCTs and what they need from them.

4.2 Research questions

The research questions define the backbone of the whole study. They are the real motivation behind it, and successful research should allow the interviewers to answer them after gathering enough data. For us, the questions we want to address are:

1. Do people find current Digital Self-Control Tools helpful? In which ways?
2. What other Digital Self-Control Tools and interventions may users find helpful?
3. Can TAP support unsatisfied user needs in DSCTs?

As we previously described, we hypothesize that the currently available DSCTs are effective in the short but not in the long term. We wanted to investigate if this is accurate and, if so,

why it happens. Although there is support for this idea based on available studies, we tried to gather a more comprehensive understanding of the user experience. Therefore, our first question seeks to explore it in general. Our starting point is to understand the users' experience and perception of the available solutions, what works, and their downsides.

The second aims to know the real user needs in the digital wellbeing field. Typically, users have a list of commercial applications available and pick the one that better fits their current needs. However, their interventions are purposely general enough to attract as many users as possible; thus, finding an application that fully satisfies their needs is rare. During the study, we tried to understand what the users would need if the available options or technologies did not limit them.

Lastly, we tried to understand if EUD, particularly TAP, could improve DSCTs. To answer this question, we gathered the user experience with rule-based interfaces or applications, or, if they did not have any, their easiness and intuitiveness perception about them.

4.3 Study structure

Contrary to quantitative studies, qualitative data rarely provides clear and not ambiguous answers. It is essential to remain objective and avoid unconsciously biasing data while defining and during the research.

For this reason, we ran four pilot tests to determine what worked and what could be improved. Thanks to them, the structure (especially for the focus groups) changed multiple times to obtain the most effective way to collect all the information of interest.

4.3.1 Questionnaire

Before the focus groups, we asked the users to compile a demographic questionnaire. Its goal is to show the user categories involved in the study, which are the ones described in 4.4.1. Furthermore, we tried to have a balanced distribution of gender and technological expertise. This effort improves the study's external validity, i.e., to collect results generalizable to the population outside the picked sample. The proposed questions were the following.

1. How old are you?
2. Where are you from? (Country level)
3. Which is your gender?
4. What is your employment status?
5. What is your familiarity with technological devices? (Likeheart scale)

In addition, we asked for the users' smartphone usage statistics for the two weeks preceding the focus group. We collected them in different periods, from the beginning of January to the middle of March. In this way, we could have a more realistic picture since each period has varying usage trends (e.g., during the holidays, people tend to use smartphones more). The idea behind this request was to contextualize the people taking part in the study, both for the research and to understand them better during the focus group. This data allowed us to interpret their answers better and to dig deeper into our follow-up questions. However, an unexpected outcome of this request was the users' reactions. Some had never checked these statistics and were astonished and even ashamed. During the study, we always tried to make the

users feel comfortable, not forcing them to send or tell anything they did not want to share, but they still felt the need to justify their usage. This fascinating aspect highlights many people's unawareness regarding their usage and dissatisfaction with it. We will dig deeper into it later on.

4.3.2 Focus Groups

Surveys, interviews, and focus groups are the most widespread qualitative techniques for collecting user data. When deciding which to adopt, we immediately discarded the first one because it did not fit well with our study. As we will see, the most valuable information collected came from the follow-up questions, i.e., spontaneous questions based on the participants' answers. Surveys are, instead, too structured to allow immediate follow-up questions.

Eventually, we decided to proceed with the focus groups. Although interviews allow us to collect more in-depth and complete information about each individual, this only partially targeted what we looked for in our research. Another critical source of information was the debate derived from each question and answer. While discussing with other people, the participants could refine their ideas, understand the motivation behind their actions, or even realize that some behaviors they consider perfectly normal are unhealthy. It would have been impossible to do the same with an interview, and it was critical to understand the problem better. Last, focus groups allowed us to host co-design sessions, which we will better describe in 4.3.3.

Although we modified the focus groups multiple times, their structure has never changed. At the beginning, after a short introduction by the interviewer, we asked some warm-up questions.

We designed them to understand participants' baseline and break the ice with them. Here, you can find the list of the asked questions.

1. Do you know there are tools to improve your phone usage habits?
2. Did you ever use them? If not, did you ever want to?
3. Do you check your phone usage statistics?
4. Do you think you are using your phone too much?
5. Do you know any tool based on the creation/personalization of rules? Did you ever use it? (possibly after an explanation about what TAP is)

The first two questions aim to know which users tried any DSCT or at least know about them. Their experience is crucial to understanding their feelings about them, improving our understanding of the first research question.

The other participants, instead, will mainly contribute to the second research question. The main motivations behind not using interventions are the unawareness about their existence, the satisfaction with their smartphone usage, or the dissatisfaction with the existing alternatives. In all the cases, it is fundamental to understand their point of view.

The third and the fourth investigate the participants' awareness concerning their usage. Finally, the last question tries to understand if any user has experience with applications implementing TAP. It is worth mentioning that almost no user knew about them, despite the existence of widely spread applications based on rules (like Alexa and IFTTT).

After this brief first section, we moved to the main questions. Their goal is to gather information about the user experience while encouraging a debate among the participants and dynamically asking follow-up questions to grasp their thoughts and needs thoroughly. In the first draft, which we adopted in the first pilot test, there were 17 of them. You can find them in Appendix A. We then decided to slim down this section because it seemed a bit dispersive, with the same topics addressed in different moments, and because for some questions (especially the ones related to TAP), the participants did not have anything to share without seeing clarifying examples.

We ended up with the following four main questions and a PowerPoint presentation constantly shown to the users during this phase. There, we summarized the questions and reported some examples of existing interfaces.

1. Tell me about your experience with tools for improving your phone usage habits. Why did you/did not you use them? What are their positive and negative sides?
2. What are your needs in the digital self-monitoring field? Do existing tools allow you to create interventions that fully satisfy your needs?
3. Do you think those tools need customization and/or context awareness? Are current solutions good enough for this? If not, what would you need/change?
4. Tell me about your experience with rule creation/personalization tools. Do you find them intuitive? Are you able to automate your actions in the way you want?

Another difference with the first version is that we avoided using any acronym or technical term to prevent user confusion.

The first question regards participants' experience with DSCTs. As we said, one of the critical aspects we wanted to understand better was the user relationship with them, considering both the positive and negative sides. This question was followed by a set of examples of interventions extracted from the most popular DSCTs so that all the users could have a panoramic view of the state of the art in this field.

The second question is more abstract. We wanted to know from the users, if they were wholly unbounded and free to imagine, what they would ideally need to improve the quality of their smartphone usage time. Of course, this led to some impossible or hard-to-implement ideas, too. For example, P17 said, "I would like an application that understands when the video I am watching is boring and proposes other videos or closes the application." Classifying a video as boring or not is a highly complex task. Nevertheless, they still helped us gain a clearer picture of their needs.

In the third question, we tried to investigate some aspects rule creation would enable in DSCTs, i.e., full customization and context awareness. With customization, we mean letting the user create something that completely fulfills their needs without adapting to externally imposed constraints. Context awareness, instead, is the capability of the tool to understand what the user is currently doing or, as the name suggests, the context around them. Based on this information, it will automatically adapt to it. This mechanism should again reduce user friction, not blocking them in situations in which they could be free or when they need to use

the smartphone to perform necessary actions. Later, we will see that the participants did not think the latter feature was useful.

The last question tried to understand people's experience with TAP interfaces and their impressions about them. Typically, almost no participant understood what we meant with rule creation tools, so it was vital to show some examples immediately. The applications we showed are Alexa (the routine creation interface) and IFTTT. The former was the first example because many people know it, and it allowed us to introduce examples based on real-life scenarios that everyone easily understood. IFTTT, instead, was selected because it is the most famous application entirely based on TAP, and it includes a marketplace of predefined rules (named applets) that are useful in showing the potential of those tools. Most users did not know them, but they were intrigued by the possibilities they enabled. P1 even decided to download and start using it after the focus group.

After verifying the efficacy of this setting in three additional pilot tests, we used it for three focus groups.

4.3.3 Co-design

HCI studies use interviews, focus groups, and surveys to collect the information they need to make informed decisions when designing for the users. In this process, users are indirectly involved, meaning their experience and answers will guide the researchers toward the best user interface possible. Still, they will never design anything in practice. Co-design tries to fill this gap. The idea is to design with the users, not just for them, by involving them in practical design sessions [35].

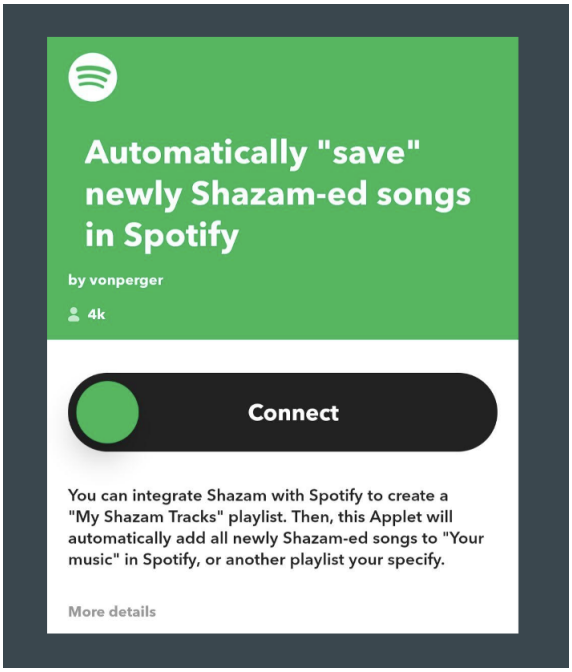


Figure 3: One of IFTTT's applets shown during the focus groups.

In our research, we used those sessions in a slightly different way. The concept remains the same (the users designed some interfaces), but the prototypes were not the primary goal. As we previously said, our objective was to gather data on users' experience with DSCTs and understand what they wanted from them. The co-design activity draws on the latter part, which is the scope of the second main question. They were particularly influential in making the participants reflect deeper on their needs, providing more detailed insights concerning what emerged during the focus group. This outcome was also possible thanks to the group division, which fostered debates and group thinking. Overall, we got a lot of inspiration for the interven-

tions people wanted, which were much more detailed than the ideas proposed during the main questions due to the more structured effort required.

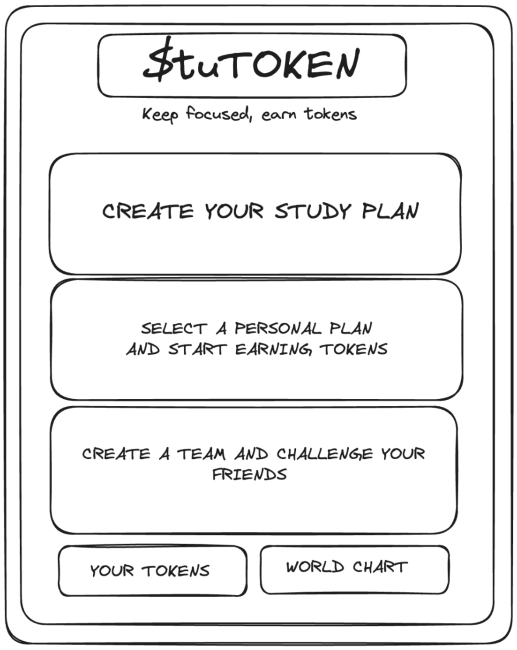
We always held the co-design session at the end of the focus group. In one pilot test, we tried to incorporate it in the middle of it. However, we immediately realized that in this way, the participants tended to repeat what they already expressed during their answers. Thus, we restored the original version.

We divided the participants into groups of two to three people. The interviewer chose the division, and it was purposely structured to create heterogeneous teams with people who could work well together. We tried to make a balance between technological familiarity, expansiveness, and ideas, and this allowed us never to have proposals that did not add any novelty. The interviewer helped the teams to follow the right path, but they never influenced in any way the participants (for example, by suggesting possible interventions) not to bias them. This aspect was critical because otherwise, the users would have sketched what they expected the interviewer wanted by them and not what they needed.

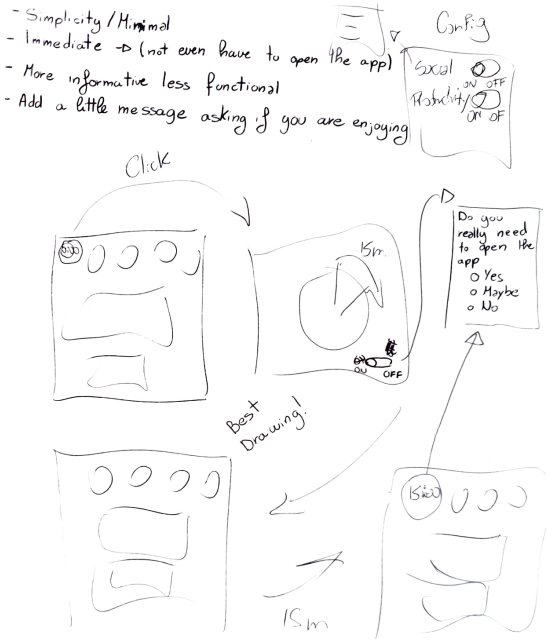
The structure of the session was the following.

1. Discuss which interventions are useful and needed in a DSCT for you.
2. Create an abstract representation of the most immediate and usable interface that would allow users to create and manage those interventions, keeping in mind what emerged during the first phase of the interview.

At the end, the teams would present their idea to the other participants and the interviewer, who could ask questions or comment about them. As we will see in 4.5, they were an invaluable source of information.



The result of an online co-design session.



The result of an in-person co-design session.

Figure 4: Two interfaces realized during co-design sessions.

4.3.4 Study

We conducted four pilot tests and three focus groups. The pilot tests were functional to refine the study structure and make the interviewer feel increasingly confident handling the conversation. They were all held in person (three in Italy, one in Chicago). For the focus groups, two were online, and one was in person in Chicago. We chose this due to the different locations of the recruited participants and the interviewer. However, although the setting changed, the content and the collected insight were equivalent to the others.

The first pilot test took place at the beginning of January. As we briefly mentioned, it included the warm-up questions, 17 main questions, and the co-design session. Although it was successful, we realized decreasing the number of main questions could improve clarity and reduce the rush in asking the questions, leaving more space for the users' answers.

In the second and third pilot tests, we tested the structure that later became the definitive one for the focus groups. We also tried to move the co-design session between the main questions instead of after them, but since we got fewer insights with this approach, we moved back to the original timeline. Furthermore, the second pilot test, held with two adults, made us understand the need for a screening questionnaire to filter out the participants who do not need and are not interested in DSCTs. We will discuss it better in 4.4.2. The final pilot test helped us test the definitive structure and the environment in which we would have held the in-person focus group.

Apart from the testing purpose, the pilot tests allowed us to gather valuable information from more participants. Many of the discussions and information that emerged during them

guided us in understanding our problem better and making more informed design decisions for the application.

Concerning the location of the study, we conducted the pilot tests in Italy in private houses. The one hosted in Chicago and the in-person focus group were in a UIC laboratory. We ran their co-design sessions using pen and paper, as seen in Figure 4. The online focus groups, instead, used Zoom. The trickiest part was to decide how to handle the online co-design sessions. All the participants were in different places, so we had to find a way of letting them collaborate profitably with as little friction as possible. We allowed them to decide whether to use pen and paper or Excalidraw, an online collaboration platform that enables users to sketch their ideas. We proposed those alternatives because we were worried some users could have difficulties using the latter option, but in the end, all the participants preferred it.

4.4 Participants

4.4.1 Target Population

One of the most critical parts of user studies is defining the target population. The people involved in the focus groups should be a representative sample of the population of interest; otherwise, the external validity will be badly affected. In our case, we discarded all the people who do not currently own a smartphone. We also did not consider older adults because, statistically, they tend not to have overuse issues with their smartphones. For those reasons, our population age target is 18 to 65. We ended up involving three main categories of people: students, young adults, and adults. As multiple studies outlined, the age category most affected by those issues

are the students and young adults, so they have been our primary focus. However, we thought collecting adults' ideas and opinions could add value to our study and design decisions.

Furthermore, we differentiated the students with an engineering background from those without it. The former category will probably have different opinions and experiences due to their higher knowledge of technology, so we tried to keep them separated to avoid an imbalance among the participants.

In the first pilot test, we included participants from all those categories. However, we realized it would have been better to separate the categories for more profitable and engaging discussions. Therefore, the third and fourth pilot tests involved only students or young adults, while the second involved only adults. For the focus groups, as we said, we gave more relevance to the students. Two included only this category (one with engineering students and one with non-engineering students), while the other included only adults.

We recruited the participants mainly from relatives and friends. We also collected some students from a UIC course on UX Research Methods to find engineering students who were not from Italy. Involving participants coming from different countries helped us improve our external validity.

4.4.2 Screening Survey

The second pilot test involved only two adults. Unfortunately for us, as it emerged from their answers, both consciously use their smartphones and are satisfied with their usage. Although we still extracted some meaningful insights from it, it is clear that the discussions that emerged

were much less profitable than the ones in all the others. Due to this, we realized the importance of having a screening survey to avoid this problem in the future.

We sent this survey to all users as the first step in recruiting. They would have continued the process only if they were eligible for the focus group (i.e., they were interested and involved in the topics we would have discussed). The questions we used to evaluate the eligibility are the following ones.

1. Do/did you study engineering?
2. How much time do you spend on average daily on your phone?
3. Do you think you are using your phone too much?
4. Are you interested in tools to monitor and/or manage your phone use?
5. Have you ever tried to use tools to monitor and/or manage your phone use? If not, would you like to try them?

We used the first question to classify the students and assign them to the proper focus group. The second allowed us to identify the people who use their smartphones too little to have overusing issues and those who use them a lot but do not consider it excessive. Furthermore, another interesting aspect is that, for this question, we did not ask for the actual number (i.e., the number provided for the usage statistics), meaning that users often reported their perceived usage instead of the actual one. Analyzing how frequently their perception matches their actual usage was a meaningful insight. The last three questions, instead, were asked to understand if

the person could be interested in what we would have discussed in the focus group. Otherwise, even if they wanted to, they would not have been able to contribute.

We consider users eligible if they meet at least one of the following conditions.

a. They think they use their phone for at least 2 hours daily.

b. They think they are using their phone too much.

c. They are interested in tools to manage their smartphone usage.

d. They either tried or would like to try tools to manage their smartphone use.

The first requirement is purposely loose because, as previously discussed, adults tend to have a lower usage time than students or young adults. Increasing this number would have made it much more difficult to find enough eligible adults, and it would not have allowed us to gather the meaningful insights we got in their focus group. The second identifies the population that would like or need DSCTs but did not know about their existence (or never tried them). The third and fourth instead represent users who want to better themselves or would like to.

The screening questionnaire was proposed only for the focus groups. Overall, it was pretty effective. From 43 responses (i.e., 43 contacted people who were available in the period of interest), we got 19 eligible participants. Some could not participate due to their unavailability in the selected time slot, but others simply were not eligible.

4.4.3 Participants' Demographic

Our study involved 39 participants overall: 20 in the pilot tests and 19 in the focus groups. Twenty identified as males, seventeen as females, one as non-binary, and one preferred not to

say. The participants ranged from 18 to 62 years old (M = 28.26, SD = 11.31). Thirty-six came from Italy, one from Palestine, one from Iran, and one from Brazil. Twenty-six were students, fourteen were workers, and one was unemployed (two participants identified as students and workers). Last, their technological familiarity (M = 3.95, SD = 0.91. One means very low, while five is very high) was at least average for all except one (who described it as low). As expected, the adults' technological familiarity (age M = 40.75, SD = 13.51. Technological familiarity M = 3.33, SD = 0.89) was lower compared to the students' (age M = 22.70, SD = 2.37. Technological familiarity M = 4.22, SD = 0.80).

For the smartphone usage (M = 278.53, SD = 109.89 minutes. Four people had the statistics disabled), it is interesting to observe if the perceived usage collected in the screening survey is similar to the actual one. For us, a prediction is correct if the perceived time is within 15% of the exact time. We considered the results of the focus groups only since, as we previously pointed out, the screening survey was not collected for the pilot tests. Furthermore, we did not consider P37 because she explicitly wrote in the answer that the value provided was the one reported by the usage statistics. Nine predictions were incorrect, and seven were correct. Indeed, sometimes the participants did not consider some usages accounted for by the statistics (P36, for example, did not include film streaming because, for him, it is passive and meaningless). Still, over half of the participants did not have a correct usage perception.

4.5 Analysis

After each focus group or pilot test, the interviewer transcribed the answers using audio recordings. To analyze the notes, we decided to use qualitative coding. In particular, we adopted

TABLE I: MOST IMPORTANT LABELS.

Label	**Meaning**
Interventions used	Interventions users use/used
Ineffective intervention	Interventions that did not work for the users (i.e., they became useless)
Negative smartphone usage	Examples of negative smartphone usage (regretted sessions)
Long-term inefficacy	Why an intervention proved to be ineffective in the long-term
Long-term efficacy	Why an intervention was effective in the long-term or what it would need to have to achieve it
Different personal needs	Examples of contrasting needs (helpful something for one user is useless for another)
Unawareness about existing interventions	Users desire an intervention that already exists (but they do not know about it)
Unawareness about smartphone usage	The users are not aware of their smartphone usage
Positive TAP perception	Positive comments or usages of TAP
Difficulty with TAP	Difficulties in using applications that implement TAP

the Reflexive Thematic Analysis, which aims to identify themes using open and organic coding [36].

During this phase, we devised many labels to grasp all the nuances of participants' answers. Ultimately, we identified the most significant themes, i.e., the labels that occurred more frequently. You can find them in Table I.

In the following subsections, we will first report some interesting interview findings. Then, we will describe why TAP would, in our opinion, improve the long-term effectiveness of DSCTs.

4.5.1 Interesting findings

One of the first things we realized is that whatever users can easily bypass is useless in the long term. Apple's option to ignore timers makes the intervention ineffective, as no one among the interviewed users respected it. For example, P26 has a timer for Instagram, which is useless for him. He rarely respects it (he seldom closes Instagram when he ends the available time). P3 also always presses the ignore button when she ends the time.

Among the interviewed users, only P2 respects the built-in timers (he uses Android, though, which has a more complex bypass strategy). P33 also used Android timers, but to ignore it, he had to change the configuration (they have no ignore button). He stopped using them since he wanted to ignore the limit and have agency over his use. Many explicitly expressed the need to block easy bypasses. There should still be a way to bypass intervention in case of solid needs, but without affecting the overall efficacy, like tokens. We will discuss it more deeply in 5.4.

We also found an increasing popularity for usage awareness. Surprisingly, some users refuse to use blocking interventions because they want to be fully in control or think they have limited effectiveness. They believe awareness is the only way to change a behavior. We expected appreciation for those interventions but did not think some users would repudiate the blocking ones completely. Interestingly, some users who expressed this view contradicted themselves shortly after. P29, for example, said that using those tools "would be a personal defeat because I want to self-limit without depending on other applications (not to give more power to the phone)." However, in another discussion, he added, "It would be interesting for social media to block when I connect to the university's WiFi. I could bypass it by turning it off, but it would

be an incentive not to use it." These contradictory thoughts result from a wrong perception of DSCTs, which should be a primary concern.

When using DSCTs, users search for something automatic and slim. Some participants pointed out that long setup or customization affects the short-term efficacy of those tools. For P33, "It is horrible when you download an application, and you have six pages of setup you must do before using it. It just makes me uninstall it and find something else." P36 agreed, saying, "We need a basic app that we install and immediately works with the minimal actions needed."

Regarding the automation, P33 used Forest, but he gave up because there were too many steps to start the application. He knows it is not many steps, but he got fed up and forgot to use it. Thus, it is clear that using DSCTs is an effort for the users, and we must make them as transparent as possible to avoid their abandonment.

Another cause of intervention abandonment is too optimistic or strict goals. P18 shared, "I created too optimistic goals and did not respect them. With less strict goals, I performed better. With strict goals, you get discouraged, and you give up." It would be better to have a smoother transition, but it is easy for the users to get enthusiastic about the novelty and overestimate themselves. Unfortunately, the stricter the goal, the higher the commitment required, often leading to the application's abandonment. DSCTs should enforce interventions but also make the users understand which ones are better for them and in which measure.

A surprising yet concerning aspect is how people normalized some highly unhealthy behaviors over time. P32 spends a lot of time on the phone, but she does something else in the

meantime. It is as if her time has "a double channel." She has many games installed because she likes to use them while watching movies. "I use it too much, but it does not take away much time." However, as P30 pointed out immediately, "multitasking is a problem, though. I also play games while watching movies, which significantly worsens the quality and quantity of attention paid to things. This multitasking is not a good use of the phone anyway, as you retain much less of what you are watching and focus less on what you are doing." She also added that "often, multitasking is rewarded. Still, it is proven that, from many points of view, it is not positive. It would be better to avoid stimulating my brain further with my phone every time I do something." Eventually, P32 acknowledged the problem. "I preach well but practice poorly because when I see children playing with an eraser while doing homework, I tell them to stop because they are losing concentration. [...] I realize it is an issue, but the big problem with social media is that they lead you to do this." The first step in solving a problem is admitting and understanding we have a problem, and DSCTs should support users in this.

Continuing to talk about the awareness issues, some users deactivated the usage statistics because they felt guilty about them. This behavior demonstrates how carefully we must design the DSCTs and their interventions. In this case, the intervention precisely hit its goal, i.e., creating a negative emotion inside the user about one of their behaviors. If not correctly calibrated, however, instead of being an aid, it becomes a bother, affecting its long-term efficacy and resulting in its abandonment. Digital wellbeing is a delicate topic, and each user is different from the others. So, customization becomes fundamental when taking into account individual

differences. As we will see in 4.5.2, this was one of the key reasons we think TAP could be beneficial.

From the screening survey, 17 out of 19 users said they use their phones too much. However, this did not always reflect in their answers, and most did not acknowledge it or, if they did, did not want to take corrective actions. P27 does not use DSCT because he would consider it a defeat. "Even though they would be convenient, I do not think I am in such a bad situation to depend on these applications." Such a conception is alarming. As we know, there are many available interventions, and each comes with a different strength and theory behind it. Even people with pretty good usage can take advantage of them. During the interview, P23, an adult, checked how much time he used WhatsApp that day. Despite having a busy day, he used it for 49 minutes (a lot for him, but not much compared to other participants). For him, notifications to understand how much time he spent on applications could be very useful. Users should see those tools precisely as P23 described them, i.e., something valuable that can support us and not as something used only by desperate people.

There is also a general awareness problem with DSCTs. Many users do not know about their existence, or if they do, they do not know that some interventions they would like to have already exist. For example, P14 said, "I did not know that there were applications that could block other applications." Similarly, P34 reported, "I use FocusTimer on my laptop. You can set limits, but changing them is not easy. Once you set the timer, there is no way to bypass it. I discovered that changing the time zone allows me to bypass it, but I rarely do that since it is hard and messes up the system. If there was something similar for my phone,

it might be better than what I currently use, but I do not know anything similar." AppBlock implements the intervention described, but she doesn't know about it. The examples in the PowerPoint presentation were helpful because most users were not aware of them nor knew something similar was possible.

Last, adults generally do not need DSCTs (or need them in a much lighter way) even if adequately picked with the screening survey. As we previously said, the screening survey helped us to filter out many participants who could not contribute to this discussion. Unsurprisingly, most of the ineligible people were adults. As we discussed, adults tend to be less affected by uncontrolled or exaggerated smartphone usage. This situation depends, of course, on their approach to this technology, which was introduced and became pervasive when they were already old enough to manage it properly. Also, they grew up in a society where they did not exist. As P24 pointed out, "smartphone usage for children and young people depends on their historical period. It was easier to play in the streets in the past than it is today." They do not wholly condemn them, but they have more instruments to maintain healthy behavior than those who grew up with them and approached them from a very young age. Also, as we said, the digital wellbeing branch is relatively new, meaning that researchers have not recognized many of the issues we are analyzing here as unfavorable since not long ago.

4.5.2 Why TAP would be helpful in DSCTs

We can now move to the last part of the study. The pilot tests, focus groups, and co-design sessions helped us answer our original research questions. We listened to participants' experiences, understanding the strengths of the available solutions and their limits, and we took

notes about what they would need. Based on that, we extracted some key insights that justify and support our hypothesis related to TAP in DSCTs.

Since the first pilot test, it emerged that all the users have unique needs, often in contrast. Apart from the fact that they may need different interventions (awareness or blocking), they typically have slightly dissimilar needs even in the same group of interventions. It frequently occurred that a participant expressed which ideal intervention they would need, and immediately after, another said it would not work for them. Unsurprisingly, the label "Different personal needs," described in Table I, was among the most popular. For instance, P39 said, "I used the iPhone default interventions. I disabled them because I got fed up with Clash Royale (a game he spent a lot of time on), so I uninstalled it and started using my phone only when needed. Uninstalling the application was helpful." However, P34 immediately replied, "When I uninstall an application, I start using the web version from the browser, so it does not work for me." In the first pilot test, P9 said, "[DSCTs] should make you feel guilty (what could you have done in this "wasted" time?)." However, P10 correctly pointed out that it might not work for some people as it might bring them down. Last, P29 said, "A friend of mine does not go to my gym with me because it does not have internet. For me, it does not make sense. Internet in the gym is unnecessary (you can download music, etc.). This behavior is symptomatic of how one panics without a phone." However, P26 disagreed. "For me, having an internet connection while doing sports is essential. I do not know in advance what music I will listen to, and I need to listen to it." Although this is not strictly related to DSCTs, it still proves that things that seem absurd to us are necessary for others because each person and their needs are unique.

Rule creation and personalization would help users create interventions that precisely match their needs without relying on external solutions that do not always fit them.

DSCT customization is needed, too. One fascinating aspect that emerged is that users may use the same application differently but still be interested in limiting it. For example, most users see social media as their most significant problem in maintaining healthy smartphone usage. P10 wants to improve her social media usage because Instagram Reels and TikTok take up much of her time. P8 agrees, defining them as rabbit holes. P11 uses Instagram more than he wants and thinks it does not add value to his life. P35 said, "I spend much time scrolling on social media. Sometimes, I open Instagram just because I am bored and start looking at all the stories without paying attention to what I see. The same happens when I am scrolling. I start scrolling compulsively, not because I see something I dislike but because I do not know what to do." Those are just some examples, but from what emerged, the use related to social media is often the most regretted one. However, they are not always poorly used. P8 used TikTok to understand a topic for an exam. She used YouTube in the past, but now she prefers TikTok because it is more immediate (it has shorter videos). She often uses Facebook to look for assignments for her students, too. P9 said, "I use Instagram often because I follow medicine pages with quick images that give me useful information. I also use YouTube a lot. I also use Facebook because, for doctors, it is handy for finding job offers." Last, P32 reported, "Instagram is also useful at a work level for me. The more you interact with a topic, the more you get similar content, and it is beneficial for inspiration for activities (especially Instagram and TikTok)." As we said, different people have different needs. Also, even people who have

found a good way of using social media would like to limit some of their usage. Still, they will look to something different compared to a user who is addicted to them and never uses them positively. TAP would enable full customization, thanks to the available triggers and actions.

As mentioned in 4.5.1, the discoverability of existing interventions is a pressing problem. Many participants in the focus groups said, "I wish something like this existed," or "I would need something like this," despite existing solutions being available. For example, P14 said, "I was unaware of the existence of these applications. Now that I know, I might use them. [...] I would need a way to force the blocks, disabling the possibility of ignoring or changing them during the day." However, some available applications allow you to do that, such as AppBlock, one of the examples in our PowerPoint presentation. After hearing about it, she was interested in it. P16 said, related to timers, "I do not feel the need for tools. Everything is up to you with the current tools, so it is useless. If there were something stricter, I would try it." Again, AppBlock is an example of what he is talking about. Thanks to the list of triggers and events, users can discover new and valuable interventions they did not know about. One could argue that it would be the same to search on the app store for the most popular keywords related to digital wellbeing and explore all the possibilities. However, almost no user does this for disparate reasons. Getting informed about new interventions by simply using an installed application they would have used anyway is far more convenient for them, and it is also more probable to happen.

One encouraging finding is that adults and students see TAP as intuitive and attractive. As previously discussed, some studies recognized TAP as the most effective EUD technique.

However, we still wanted to ensure our participants shared this perception. Notably, when asked for opinions about that, all the negative comments we received were related to some applications and not to TAP itself. Most regarded home automation, with Alexa being our most famous example. However, they were positive when we asked the users to focus on the rule-creation process.

The most concerning and interesting population was adults due to their lower technological familiarity in general, and they were enthusiastic. P23 said, "I am very intrigued by rules. [...] I would be very interested in TAP, not related to home automation, too." Similarly, P22 reported, "I find rules useful. I do not see an immediate example, but I see it positively regarding simplifying/using technology." Furthermore, as P24 pointed out, "Creating rules would be feasible; it takes some time to make them, but you find the way. In case of difficulties, there are external aids (internet, etc.)." Those feelings were shared by students, too. P17 said, discussing his experience with Alexa's routines, "The rule creation part was easy. Still a little bit confusing because there is a lot of stuff, but when you understand it, it is easy." P36 thinks, "they are surely useful. I never thought about them. They are also pretty intuitive."

The users proposed many ideal interventions during the focus groups and the co-design session, which we will better analyze in 5.4. It is impossible to create one app or one functionality for each, but it is possible to let users create them by defining rules. Combining triggers and actions in different ways unlocks new possibilities, satisfying more users.

The reactions to context awareness slightly surprised us. Thanks to TAP, it would have been straightforward to incorporate it in a DSCT. There are working examples of this, especially in

the IoT field, so we expected users to appreciate it. However, we received mixed feedback, with most participants not wanting it or considering it harmful. As P28 shared, "I prefer to have control over when to activate the blocks so that I am not even more at the phone's service. Having control requires a lot of self-discipline (you must not bypass the blocks/controls)." One of the most exciting aspects for us was the possibility of automatically removing the interventions when they were not needed (i.e., in moments in which they would have caused friction in the user) without affecting the overall effectiveness. However, P14 said, "If my plans change, it would not be a problem because I use my phone often at home and not when I am out. Having some applications blocked in those moments would not cause discomfort (especially for social media)." Similarly, P30 would not need it because she has a fixed routine.

There were some positive comments, too. P18, for example, shared, "My smartphone automatically turns on the Do Not Disturb mode based on the context. I would like to have the same with DSCTs (for example, Instagram disabled during lectures)." However, since most users were not enthusiastic, we decided not to include it in the final application.

To conclude, since TAP enables full customization and adapts to changing user needs, it could improve the long-term efficacy of DSCTs if adequately designed. As we analyzed, the leading causes of abandonment are friction and ease of bypassing the intervention. Long-term efficacy requires total commitment, and the possibility of having interventions that perfectly fit your needs will help it. Furthermore, it becomes your choice what to activate (it is not just downloading an application anymore, but also creating the rules), which could make you want to follow it more strictly. This approach would require, in fact, two commitments from the

user: downloading the application and creating the interventions. Each intervention is related to a higher level of consciousness, thus making the user more willing to respect it. If, at any time, the user feels the intervention is less effective, they will be able to modify or change it instead of altogether abandoning the DSCT.

CHAPTER 5

HABIT ALTERATION MODEL

A theoretical framework is needed to provide a theory to support our observations and to make informed design decisions. Indeed, our study provided much information to help us understand how to improve the existing DSCTs. Still, behavioral theories describe the most effective way to enforce a change in the users.

After another minor literature review phase, we understood the Habit Alteration Model [37] was suitable for us. In this chapter, we will describe it, see how it supports our findings, and provide a mapping between it and the desired interventions shared by the participants. Last, we will review some of its design principles that helped us find a better design for our application.

5.1 Overview

Habits are central to this theory. They are "learned impulses to perform a particular behaviour, triggered outside of conscious awareness by a particular context. Habitual behaviour is learned behaviour that is 'frequently repeated, has acquired a high degree of automaticity, and is cued in stable contexts [38]'. [...] In contrast to common usage of the word habit, we do not define it as the behaviour itself. Instead, following Gardner [39], we define a habit as a link represented in associative memory between a certain context and a specific response." From this definition, it is clear that we can classify many smartphone usage patterns as habits. One

of the most frequent examples is social media. For them, typically, the context is being bored while doing something that can be interrupted, such as studying or having some free time. The automatic action is using social media, often trapped by some dark patterns such as Instagram Reels or TikTok. For example, P3 said, "I was at work and did not have much to do, so I used TikTok for 2 hours. In the end, I regretted it because I could have studied."

The idea behind the HAM is to break bad habits by establishing new, healthier ones. With bad habits, we refer to the habits people regret. One example is using their smartphone too much, but others could be eating too many snacks during the day or smoking. The HAM is, in fact, a general-purpose framework that works in all the contexts for which we have formed some habits, including digital wellbeing. A crucial intuition is that habits are the default behavior when people cannot or do not want to make effortful decisions about their behavior [40]. So, if we can break the old bad habits with new positive ones, we can introduce changes that will likely persist in the long term. According to them, this is why most Digital Behaviour Change Interventions tend to be just temporary alterations of the user's behavior: their most common technique is to provide information, not create new habits.

5.2 Relevant theories

The HAM is based on multiple behavioral theories. In this section, we will analyze a few of them that are particularly relevant to our study.

5.2.1 Behaviorism

Behaviorism is a theory that focuses only on observable actions to explain behaviors since we cannot rigorously observe cognitive constructs. According to it, habits are stimulus-responses

pairs formed through two mechanisms of associate learning. The first one, classical conditioning, is the pairing between stimulus and responses. The second is operant conditioning, which associates each pairing with a positive or negative outcome. Rewards should strengthen this link and make it more likely to happen. When rewards are not needed anymore, i.e., removing them does not alter the behavior, we can consider the habit as established. The most effective way of proceeding is to give rewards with a fixed rate, but not always at the same given time.

The biggest criticisms of behaviorism are related to the complete ignorance of cognitive constructs. Due to this, it cannot explain the influence of higher-level concepts such as goals and outcome expectations. It also does not consider the role of cognitive constructs like mood, which can alter how a person behaves.

In our study, examples of positive and negative behaviorism emerged. For example, some participants reported that having real-life or virtual rewards would incentivize them to respect the interventions. As P30 said, "The problem is that these applications (DSTCs, editor's note) do not encourage you much to avoid this behavior. For example, if I knew that on Forest, every plant planted in my forest is also planted in real life, doing something positive for the environment, I would be more motivated to use it seriously. Having tangible benefits, not just virtual ones would be more satisfying." This suggestion is a clear example of positive behaviorism.

Concerning negative behaviorism, some participants suggested an intervention that, under certain conditions, worsens the user experience. Possible examples were increasing the number

of advertisements in the application, making the phone vibrate, or reducing internet speed to increase the content loading time.

The Modern Habit Theory, one of the three theories upon which HAM is based, integrates behaviorism with goal-directed cognitive-reasoned action theories [41].

5.2.2 Dual Process Theory

The dual process theory is one of the most fundamental theories in digital wellbeing and will be the base for the application's design. It states two distinct processes behind behaviors: Type 1 (automatic) and Type 2 (conscious). Habits are, of course, part of the Type 1 processes. Our behaviors are the result of the interplay between them.

Type 1 processes become more relevant as Type 2 resources are depleted or in specific contexts, such as time pressure, during distractions, or with a high cognitive load. Furthermore, the predominance of one of the two processes highly depends on the personality, too. For those reasons, their influence on the behavior will vary over time and between people.

The most crucial difference with Behaviorism is that, for this theory, habits rest on cognitive constructs. So, we can alter them through both cognitive and behavioral techniques.

As we will see in Chapter 6, one of the first intervention classifications we presented (awareness and blocks) is strictly related to this theory.

5.2.3 Goal Setting Theory

The Goal Setting Theory explains how to form goals that encourage behavioral repetition when Type 2 processes dominate over Type 1. To be effective, the user must accept the goals, progressive feedback is fundamental, and proper calibration is required. With calibration, we

refer to the goal difficulty and specificity. Complex and specific goals are more effective than vague and easy ones.

This theory's biggest problem is measuring goal commitment and difficulty. The best way to proceed is to take into account individual differences.

5.3 Model

The Habit Alteration Model synthesizes the Modern Habit Theory, the Dual Process Theory, and the Goal Setting Theory. It is a practical conceptual model that can be readily applied to DBCIs. Its goal is to provide a graphical simplification (visible in Figure 5) of how, according to those theories, external and internal factors contribute to generating habitual and non-habitual behaviors.

It all starts from the context, which includes cues. Type 1 processes map those cues into impulses, and Type 2 processes into intentions. A response is generated based on their relevance (which depends on the specific individual) and strength. The response causes a behavior. Finally, the response and the outcome caused by the behavior feed back into the model.

The solid lines in Figure 5 represent always-running processes, while the dotted ones may run. With repetitions of the same behavior in stable contexts, those cycles become more automated, i.e., the behavior slowly moves from a Type 2 process to a Type 1. This mechanism is the habit formation process. The habit disruption, instead, relies on the Type 2 processes overriding the automated Type 1 response, given that we have enough Type 2 resources to enforce it.

We will now dig deeper into each phase.

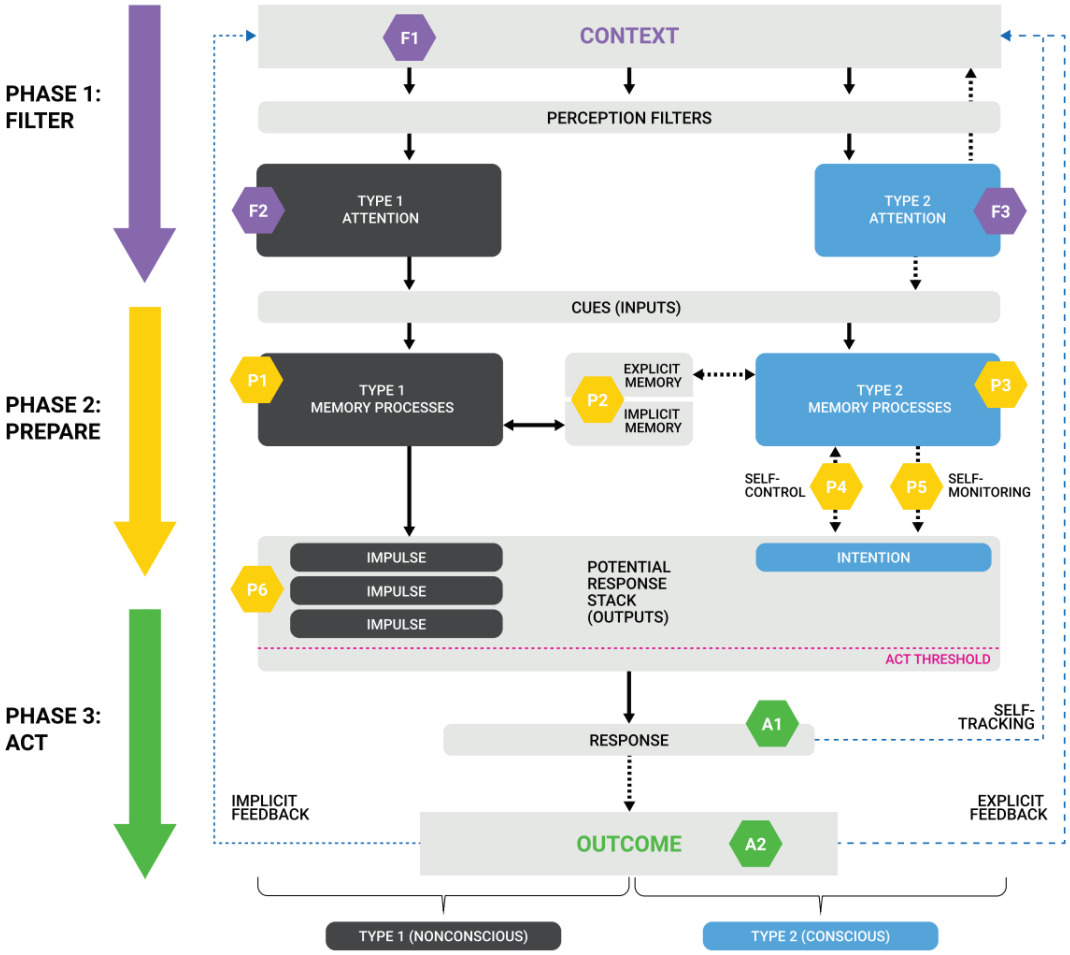


Figure 5: A visual representation of the HAM.

5.3.1 Filter

In the filter stage, we extract from the context the set of cues that will be relevant to our behavior choice. Cues can be external (physical objects or people) or internal (mood or physiological drives like hunger) features. From this set, the user will filter out the features

based on implicit (Type 1) and explicit (Type 2, i.e., directed attention) attention processes. Of course, the intervention and relevance of this latter process depend on the availability of Type 2 resources. The result of this phase is the cues, which are the input for 5.3.2.

5.3.2 Prepare

Type 1 processes extract impulses from the cues, while Type 2 intentions. The impulses and intentions compete to become the behavior and are represented in a *Potential Response stack*. Type 1 processes generate multiple impulses from the set of cues fast and in parallel. Of course, not all impulses are habits. Type 2 processes, instead, derive from explicit user goals via self-control and self-monitoring. Typically, self-control intentions are intentions not to act, while self-monitoring ones are intentions in line with our goals.

The order in the stack, which depends upon multiple factors, reflects the strength of the intention or impulse: the higher, the more frequently that behavior previously occurred.

5.3.3 Act

The behavior to be performed depends on the status of the Potential Response stack. If one response (an impulse or an intention) crosses the act threshold, we will perform it. In case of conflicts between different or within the same response type, Type 2 processes will intervene to decide which to perform.

Once we enact the behavior, there may be an outcome. The information about it will feed back into the model and impact the following cycles.

Type 2 processes are crucial to breaking unwanted habits. To do so, they populate the stack with intentions to compete with unwanted impulses and enter the arbitration phase in case of

conflicts. However, if Type 2 resources are unavailable, the most likely response is the highest impulse in the stack. Impulses appear more quickly, making them more likely to reach the act threshold.

5.4 Desired Interventions

The HAM proposes some strategies to modify people's habitual behaviors. After analyzing all of them, we extracted the ones that were more relevant to what emerged in our study.

- **Alter Context (Filter)**. This strategy adds or removes cues to alter the responses inserted in the stack. Another possibility is to alter the user's mood to make them behave in a specific way.
- **Instinctive paths (Filter)**. The idea is to exploit the automatic context-response paths within humans to enforce a behavior change. Possible examples of those paths are the influence of auditory or social priming, i.e., the fact that people tend to repeat the behaviors seen by others.
- **Implementation Intentions (Prepare)**. They are specific *if-then* plans that link a context (if) with a wanted response (then). By leaving the control to the selected *if*, they aim at automating the response. They work by increasing the accessibility so that the desired behavior is higher in the response stack.
- **Provide Information (Prepare)**. We provide information to the user that should change their decisional balance. This strategy assumes a rational choice model, i.e., the person will change behavior in light of the new information.

- **Just-in-Time Reminders (Prepare)**. Based on the context, we can propose to the user just-in-time reminders to behave in a specific way. They can help the user to form new habits and break old ones. This strategy may, however, make the technology that provides the reminders become a cue itself, thus making the habit less resistant.
- **Train Self-control (Prepare)**. Self-control is the ability to "alter [your] own behavioral patterns so as to prevent or inhibit [the] dominant response [42]." It is a Type 2 mechanism that allows us to resist impulses to behave unwantedly. It should become a Type 1 process through repetition to be effective in the long term. It is mostly useful for breaking unwanted habits but can also be used to form new ones.
- **Self-monitoring (Act)**. Using information from self-tracking, the person can form new intentions. This strategy works because the Type 2 systems get to know the consequences of their automatic Type 1 responses, which previously were unknown. It can motivate to enforce a change.
- **Revalue Outcome (Act)**. The idea is to provide rewards for good behaviors and punishments for bad ones. Rewards are unnecessary if the behavior is intrinsically rewarding, but they may speed up the process. This strategy works because rewards increase the expected value of a response, i.e., they move it up in the stack. Punishments are, instead, rarer. A common idea related to this strategy is gamification.

We then mapped all the desired interventions proposed by the participants by using those strategies. This way, we proved that a behavioral theory can support users' desires. If they are

adopted and followed with dedication, they can indeed form new healthier habits, improving digital wellbeing.

Here, we can find the list of desired interventions, their mapping into rules, and the HAM principle behind them. We split the table into two parts. The first reports the desired interventions emerged during the focus groups, and the second during the co-design sessions.

TABLE II: DESIRED INTERVENTIONS.

Intervention	**Trigger**	**Action**	**HAM**	**Note**	**Requests**
An application timer without the option to ignore or change it during the day.	Usage time of an application or time of the day.	Block the application.	Train self-control.	The rule lock avoids modifications throughout the day.	8
Block app reinstallation.	Application installed on the smartphone.	Block the application.	Alter context.	It is probably impossible to block the installation, so the easiest way is to block it forever.	1
Intervention rotation.	Not relevant.	Not relevant.	Nothing specific.	This can be achieved with automatic rule creation since, in the interview, the user referred to something automatic. So, the trigger-action mapping here is useless since it is more of an internal mechanism.	1

Table II – continued from previous page

Intervention	Trigger	Action	HAM	Note	Requests
Interventions that make you feel guilty, showing you what you could have done instead of using the smartphone.	Application closed or usage time.	Show a message.	Just-in-time reminders.		1
A comparison with the average use or use considered healthy.	Not relevant.	Not relevant.	Provide information, Instinctive paths.		1
An intervention that does not allow you to open the application.	Application launched.	Block the application.	Train self-control.		1
Social media to implement controls themselves.	Not relevant.	Not relevant.	Alter context.		1
Be forced to watch each TikTok for at least a few seconds before being able to scroll.	Application currently opened.	Delay the swipe gesture.	Alter context.		1
Have more advertisements to be less inclined to use the application.	Application currently opened.	Insert pauses.	Alter context.	It is unfeasible to add advertisements to other applications. However, we could mock this behavior by manually pausing the application for N seconds with a black screen, for example.	1
Remove anything related to profiling, like suggested posts.	Not relevant.	Not relevant.	Alter context.		1

Table II – continued from previous page

Intervention	Trigger	Action	HAM	Note	Requests
Limit what others do on the phone regarding privacy.	Not relevant.	Not relevant.	Not relevant.	It is not strictly a digital wellbeing intervention.	1
Standard blocking, but not to be able to modify it when studying.	Time of the day.	Block the application.	Train self-control, Implementation intentions.	The user can lock the rule at the beginning of the study session.	1
Precise scheduling of blocks (blocking at specific times) to program the day.	Time of the day.	Block the application.	Train self-control.		1
Block the phone completely, apart from emergency functions, for a fixed amount of time (like a timer).	Not relevant.	Not relevant.	Implementation intentions.	Rule mapping is unnecessary since the user refers to a block manually activated.	2
Real-life rewards in DSCTs.	Not needed.	Not needed.	Revalue outcome.	Rule mapping is unnecessary since the rewards derive from respecting or not following the rules.	1
Stop having stories on Instagram, but still see posts.	Not needed.	Not needed.	Alter context.	Feature-specific interventions require more ad-hoc mechanisms than general rules.	1
Deactivate YouTube shorts.	Not needed.	Not needed.	Alter context.	Feature-specific interventions require more ad-hoc mechanisms than general rules.	1

Table II – continued from previous page

Intervention	Trigger	Action	HAM	Note	Requests
A timer that allows you to end your activity (even if you finish your available time), and then you are cut off when you exit.	Usage time of the application.	Block the application.	Train self-control.	It can become a timer option.	1
Usage reminders (like "10 minutes left") with timers.	Usage time of the application.	Show a message.	Just-in-time reminders.		2
Get notifications that tell you how much you have been using the application or the smartphone that day.	Usage time of the application.	Show a message.	Self-monitoring.		8
Customize the YouTube client.	Not needed.	Not needed.	Alter context.	Feature-specific interventions require more ad-hoc mechanisms than general rules.	1
An intervention that understands when the video is boring (for example, when the video is picture-in-picture) and proposes other videos or closes the application.	Video in picture-in-picture for N minutes.	Show a message or close the application.	Just-in-time reminders, Train self-control.	Here, the rule mapping is complex since *boring* is a subjective concept. We propose a mapping based on what the user said during the focus group.	1
An intervention that asks if you are interested in the application you are using or if you would rather close it.	Complex.	Complex.	Just-in-time reminders.	Again, the rule mapping is complex since *interesting* is subjective.	1

Table II – continued from previous page

Intervention	Trigger	Action	HAM	Note	Requests
Something easily reachable, like a dedicated app or a widget.	Not needed.	Not needed.	Self-monitoring.		1
A filtering mechanism for Instagram posts.	Application currently opened.	Blur/filter out content with a specific topic.	Alter context.	The topic could be a rule option. Again, this is probably unfeasible with a general mechanism such as rules since it is highly implementation-dependent.	1
Functionalities to gently guide the user toward his usage goal.	Not needed.	Not needed.	Train self-control.		2
Something to make the user understand whether they are in control of their smartphone usage.	Not needed.	Not needed.	Self-monitoring.		1
A mechanism to block continuous scrolling.	N minutes scrolling passed.	Disable scrolling or close the application.	Alter context.	Disabling scrolling may not be feasible.	1
Limit some application features.	Not needed.	Not needed.	Train self-control.	This is probably unfeasible with a general mechanism such as rules since it is highly implementation-dependent.	2

Table II – continued from previous page

Intervention	Trigger	Action	HAM	Note	Requests
Limit passive phone and social media use but allow users to choose for active use.	Not needed.	Not needed.	Train self-control or self-monitoring (depending on how it's achieved).	Rule mapping is unnecessary since it can be achieved using rules in general (but no one in particular).	1
Progressive timers (a certain amount of time in the morning, afternoon, etc.) rather than daily ones.	Time of the day and usage time of the application.	Block the application.	Train self-control.		1
An intervention that profiles your usage and habits and automatically proposes a study plan.	Not needed.	Not needed.	Train self-control.	Rule mapping is unnecessary. It's an automated mechanism that analyzes your behavior and creates rules automatically. Rules are involved, but again, no one in particular.	1
An intervention that allows you to use an application each day for a certain number of hours for free, with the following ones available behind a paywall.	Usage time of an application.	Block it.	Train self-control.	There should be a rule option to allow the user to pay to overcome the limit.	1

Table II – continued from previous page

Intervention	Trigger	Action	HAM	Note	Requests
Make the user notice when they open a distracting application passively after performing a useful action.	Not needed.	Not needed.	Self-monitoring.	Rule mapping is unnecessary since, in this case, the user asked only for statistics and not for an intervention such as a popup message.	1
A mechanism to avoid unlocking the smartphone to perform a useful action and doing something completely unrelated instead.	Smartphone unlocked.	Popup message to prompt the user to their goal for the usage session.	Train self-control.		1
An intervention that immediately works, with little to no customization required.	Not needed.	Not needed.	Not relevant.		1
A token-based timer. You can override the timer, but only a limited number of times.	Usage time of an application.	Block it.	Train self-control.	There should be an option to allow the user to configure the token mechanism.	1
An intervention that purposely worsens the user experience.	Time of the day or usage time of an application.	Worsen the experience.	Alter context.		1
_____	_____	_____	_____	_____	_____

Table II – continued from previous page

Intervention	Trigger	Action	HAM	Note	Requests
When the user opens a distracting application, they decide how long to use it. One minute before the timer ends, the application turns grayscale. At the end of the timer, the application locks.	Application launched.	Popup message to prompt the user to the desired usage time.	Train self-control.		1
Monitored applications consume more battery than necessary, so if the user must go out, they will not use distracting applications too much.	Time of the day and application currently opened.	Increase battery consumption.	Train self-control.		1
When the daily timer ends, the application disappears from the home screen.	Usage time of the application.	Remove the application from the home screen.	Alter context.		1
Increase the number of advertisements in distracting applications.	Already discussed.	Already discussed.	Already discussed.		1
You can keep using the application at the end of the timer, but the phone vibrates (usage becomes annoying).	Usage time of the application.	Make the phone vibrate.	Alter context.		1
When you open a distracting application, it turns black and white.	Application launched.	Turn the screen grayscale.	Alter context.		1
A division between necessary and playful apps.	Not needed.	Not needed.	Not relevant.		1

Table II – continued from previous page

Intervention	Trigger	Action	HAM	Note	Requests
Intentionally make the interface of specific applications ugly.	Not needed.	Not needed.	Alter context.	The rule mapping is unnecessary since it is implementation-dependent.	1
A notification that asks if you are sure you want to open an application.	Application launched.	Popup message to prompt the user if they want to use it.	Just-in-time reminders.		1
Disable swipe down in Reels/TikTok and change the gesture (explicitly exit the video and open another one).	Already discussed.	Already discussed.	Already discussed.		1
Calendar-style interface. It shows every day, each with all the time slots. You can decide what to do in each time slot (study, etc.) and choose which applications to block for each slot. You can allow, block, or partially block applications (e.g., allow messages from specific contacts).	Not needed.	Not needed.	Train self-control.	The rule mapping is unnecessary (rules cannot implement the described calendar interface). However, rules based on the time of the day are feasible.	1
Gamification. You get points for each successful task (e.g., timer respected); at the end of the week, you get rewards based on the points.	Already discussed.	Already discussed.	Already discussed.		2

Table II – continued from previous page

Intervention	Trigger	Action	HAM	Note	Requests
You receive notifications when you use specific applications while you have an event on the calendar.	Event on the calendar and application currently opened.	Show a message.	Just-in-time reminders.		1
Automatic shutdown of notifications/ringtones.	Time of the day.	Disable notifications or turn off the ringtone.	Alter context.		1
Limit unnecessary applications at specific times, but keep active the channels you want to be reachable on (specific contacts, etc.).	Already discussed.	Already discussed.	Already discussed.	The active channels could be implemented as an option, although they would again be implementation-dependent (so rules may not be ideal).	1
Form teams to challenge with friends (also with weekly rankings).	Not needed.	Not needed.	Instinctive paths.		1
Group study sessions, in which the distracting applications are blocked for all the group members.	Not needed.	Not needed.	Instinctive paths.		1
Allow users to filter notifications (permit calls/messages only from specific contacts or allow only certain types of notifications) or mute them entirely.	Already discussed.	Already discussed.	Already discussed.		1
Track the daily number of phone unlocks.	Not needed.	Not needed.	Self-monitoring.		1

Table II – continued from previous page

Intervention	Trigger	Action	HAM	Note	Requests
Token-based mechanism to bypass timer. Once skips are exhausted, specific actions (to be performed in real life) can still bypass the block.	Already discussed.	Already discussed.	Already discussed.	The bypass can be a timer option.	1
When an application opens or a notification telling you how much time you used the application that day shows up, a question appears: whether you want to use the app or how you're feeling about your usage.	Application launched or usage time of the application.	Popup message.	Just-in-time reminders.		1
You receive a popup where you can set timers for that day the first time you unlock your smartphone in the morning. You can customize the limits for each application in a specific list. You can also set how many times you can ignore the limits.	Not needed.	Not needed.	Train self-control.	Rules can implement the locks, but this intervention doesn't refer to any rule in particular.	1
Create rules that determine the consequence of bypassing the timer.	Usage time of the application.	Not specified.	Revalue outcome.		1

Table II – continued from previous page

Intervention	Trigger	Action	HAM	Note	Requests
Depending on your focus mode, you can change the home screen layout.	Not needed.	Not needed.	Alter context.	The rule mapping is unnecessary (rules cannot implement focus modes). However, it's possible to create a rule that changes the home screen layout based on the time of the day.	1
Token-based system for timers. Each token grants a fixed amount of additional time on that application. Once the tokens are over, you can still use the applications, but their performance will always get worse.	Already discussed.	Already discussed.	Already discussed.		1

5.5 Design Principles

The HAM proposes some design guidelines to create more effective DBCIs. We tried to follow them when deciding on the application layout.

The first suggestion is to simplify the target response, i.e., the actions. Also, the cues (the triggers for us) that compose a stable context should be as restricted as possible. Both strategies allow a reduction of the time needed to reach behavior automation. We can achieve it with rules by extracting a list of primary triggers and actions from participants' desired

interventions. They typically are essential, such as time of the day or block the application. If they need to be more complex, we integrated some options into them but tried to keep them minimal.

Type 1 and 2 behaviors influence individuals in different ways. For this reason, other people may need different intervention types and strengths. Rules perfectly accomplish this because they allow full customization of the interventions. We divided Type 1 and 2 interventions, too, because some users may want only to use one of those categories, either because it is more effective for them or because of personal preference. Furthermore, the interventions should dynamically adapt and change to avoid technological dependence and reactance. For this reason, we introduced an automated rule-creation mechanism.

HAM also proves that behavior depends on Type 1 and Type 2 processes. Thus, DBCIs should not focus just on one of them because otherwise, they will be less effective in changing user behavior. For this reason, we included rules for both process types in our application.

DBCIs should be designed for persistence, i.e., making the user as little dependent on them as possible. They can be defined as successful if, after some time, the user does not need them anymore to enforce the desired behavior. Again, thanks to automatic rules and their dynamic adaptation, we can continuously monitor user progress and adapt to make them increasingly independent.

As we previously discussed, interventions may provoke reactance in the user. Reactance is people's reaction when they feel their freedom is restricted and try to regain it. This feeling is one of the most frequent motivations behind DBCIs (and, more specifically, DSCTs) aban-

donment. For this reason, blocking interventions should be proposed only when appropriate. In our application, we delegate this aspect to the user. Since the user has defined the rule causing the annoying behavior, we trust they will feel committed to it, thus not giving up on the intervention. If, at any moment, they think it is too strict, they can always edit or delete it while keeping the other interventions. Concerning automatic rules, we decided to allow the user to edit or delete them to decrease the reactance caused by them. Furthermore, they can be disabled at any time (by default, they are not active).

Last, the proposed interventions should be ethically designed, i.e., the user should always be aware of what the intervention is doing and why. Of course, this aspect is trivial with TAP since the user decides which rules to enforce. We show automatic rules alongside user-defined rules so they are fully transparent. They also need to be manually turned on since, by default, they are off. For this reason, the user will be fully in control of them. We preserve privacy, too, because all the data the application uses never leaves the user's device.

CHAPTER 6

APP IMPLEMENTATION

The last part of the thesis consisted of designing and developing a prototype of a TAP-based DSCT. Up to now, we analyzed the problem mainly from a theoretical point of view. In Chapter 2, we went through the developments in the digital wellbeing field in the last few years and the criticalities of the existing DSCTs. In the literature review phase, we analyzed some studies that tried to improve the effectiveness of those tools, others that proposed some design guidelines, and some about EUD and, more specifically, TAP. Inspired by them, we designed and ran a user study to analyze what is not working in DSCTs long-term, what the users need for interventions, and if TAP was intuitive and understandable. The gathered insights confirmed most of our hypothesis, and they were supported by the Habit Alteration Model, which provided a theory that better explains the participants' answers and gave us some suggestions about how to make the desired behaviors become habits.

Providing something practical based on everything we understood can add value to our research. Moved by this idea, we designed and developed an Android DSCT based on rules.

6.1 Why an Android application?

Digital wellbeing is mainly related to smartphone usage. This focus is due to the considerable smartphone improvements over the last few years. Today, they allow users to do most of the tasks they must perform throughout the day. Laptops are still better in some aspects, such as

writing comfort or computational power, but the big difference with smartphones is that they are not as easily accessible.

Since the beginning, we focused our attention on them. Almost all the DSCTs available on those devices are applications. Other solutions, such as browser extensions, are not as effective because they allow us to intervene in a limited portion of the usage. As emerged during the interviews, however, the most pressing issues for users are social media and the easiness of getting distracted. An application allows us to monitor and improve the overall usage, not limited to a specific app or category.

We then had to decide the platform where we wanted to develop it. The available alternatives were Android, iOS, or a cross-platform application. We immediately discarded the last option. Interventions are, in fact, different by standard app functionalities. They often require higher-level privileges to perform actions such as getting usage statistics and closing or showing messages over other applications. For this reason, developing a solution that would work on both platforms would have been cumbersome - if not unfeasible.

We eventually selected Android for two reasons. First, as of April 2024, Android is by far the most widespread mobile Operative System, with a market share of 70.9% [43]. Second, iOS is known to be a more closed environment, with higher limitations in what an application has access to or can do. Android is, instead, less strict, thus giving us more freedom to implement the desired actions.

6.2 Design

Based on the user study results, we devised three initial design proposals. While they all involve rules, they differ in the layout or the proposed features.

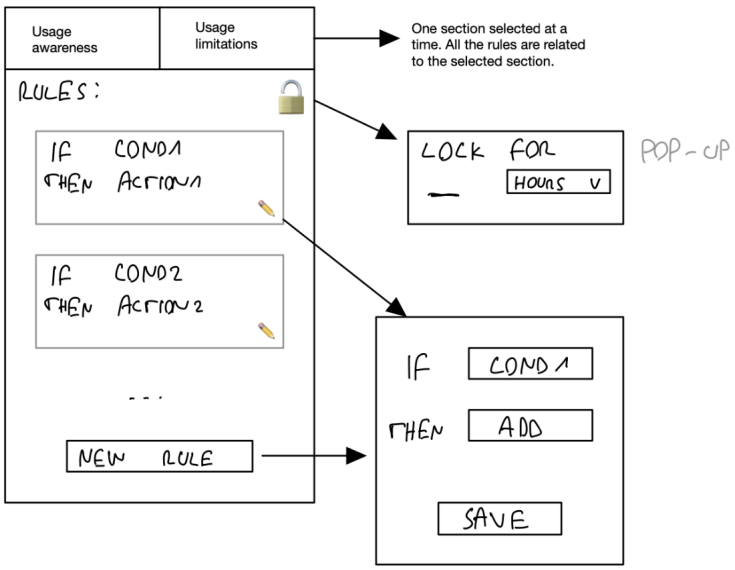


Figure 6: The first design proposal.

The first design proposal, visible in Figure 6, is shaped upon the Dual Process Theory. We split the interface into two sections- one related to Type 1 (usage limitations) and one to Type 2 (usage awareness) processes. We consider the awareness interventions to belong to Type 2 processes because they do not impede the user in any way. They may remind them about their

goal or make them aware of their usage, but it is the user with their Type 2 processes who must consciously decide to stop using the smartphone. Usage limitations are, instead, considered Type 1 interventions because they act on behalf of the user, impeding the unwanted automatic Type 1 processes.

Based on what emerged from the interviews, some users only want something to increase their awareness, while others think this wouldn't be enough for them and thus ask for something more substantial. This solution allows them to build the interventions that best fit their needs, using one or both sections.

The rule section is based on TAP. Users can create, see, or edit their rules. There's also a lock, which blocks any modifications for a certain period. Its usage is entirely optional. We included it because people who said it would be strongly needed in the interviews can take advantage of it. People who think it's useless or harmful can ignore it.

The rule creation or editing page is essential not to overwhelm the user. The available triggers and actions list will depend on which section contains the rule.

The second proposal (Figure 7) was inspired by what emerged in the interviews and the HAM. Like a mentor, the idea is to have an application that automatically learns and adjusts interventions based on the user's behavior and usage. Some rules are automatically created/edited, and the user can edit or create new ones.

The solution on the left is the most essential one. It includes a list of all the active rules, highlighting the ones created by the user. This arbitrary choice can be changed by highlighting the rules automatically created or marking each with its creator. Instead, the one on the right

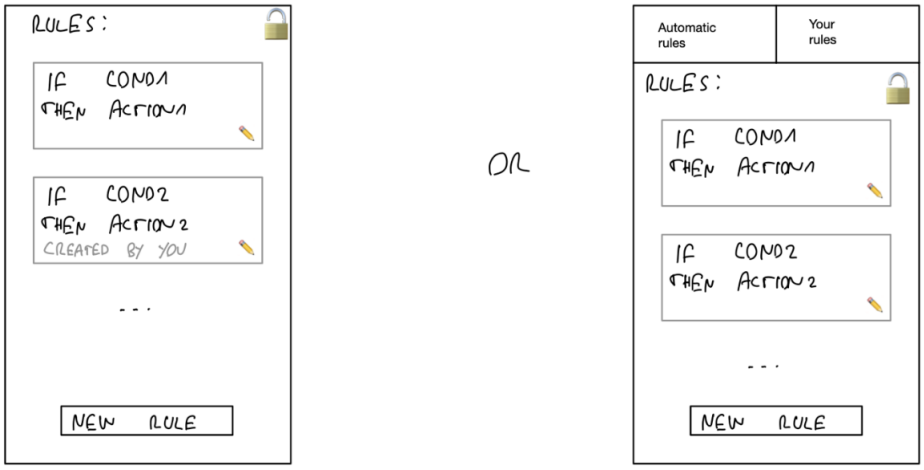


Figure 7: The second design proposal.

has two different pages, one for the automatically created rules and one for the user-created rules. Of course, the option to add a new rule is shown only on the latter screen (it wouldn't make sense to include it in the former one, too).

We kept some choices (like the lock and the rule creation/edit interface) from the first proposal.

Last, the proposal in Figure 8 is similar to the first. The main difference is in how we arranged the rules. The first screen will show a list of intervention categories. When the user clicks on a category, he navigates to the rule page. Here, all the rules shown will be related to the selected category. The rule creation/modification page will automatically adjust the list of triggers and actions based on the category.

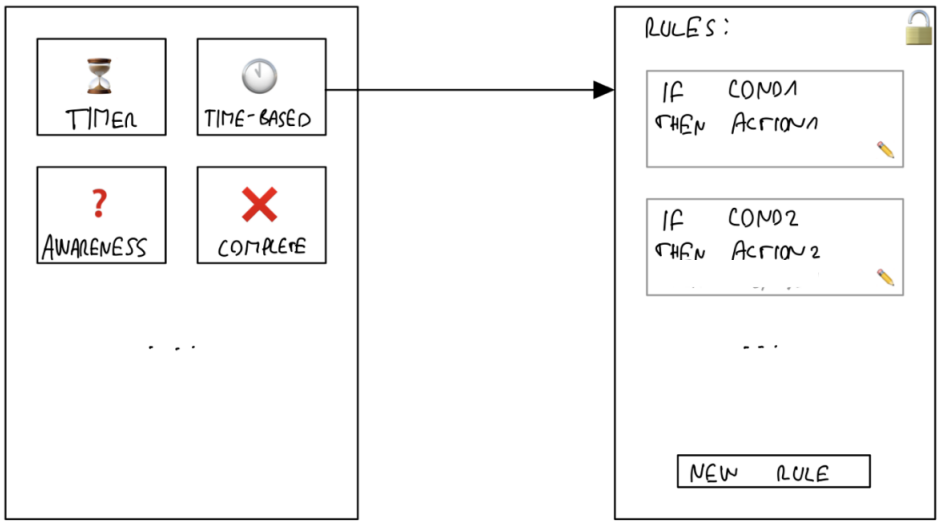


Figure 8: The third design proposal.

After brainstorming, we agreed the most exciting alternatives were the first two. In particular, we appreciated the division between awareness and blocking interventions since, as it emerged from the interviews, some users were entirely against the latter intervention type and probably would not even consider a DSCT without this division. However, the automated rule management was interesting, too. Not only was it one of the design principles derived from the HAM, but it also emerged during the interviews that there was a need for a mechanism to support users during the selection and configuration of interventions.

Ultimately, we merged them in an application storyboard, visible in Figure 9. We introduced a small setup to implement the automatic rule functionality considering user preferences run the first time the application opens. In it, we ask what type of interventions the user is interested

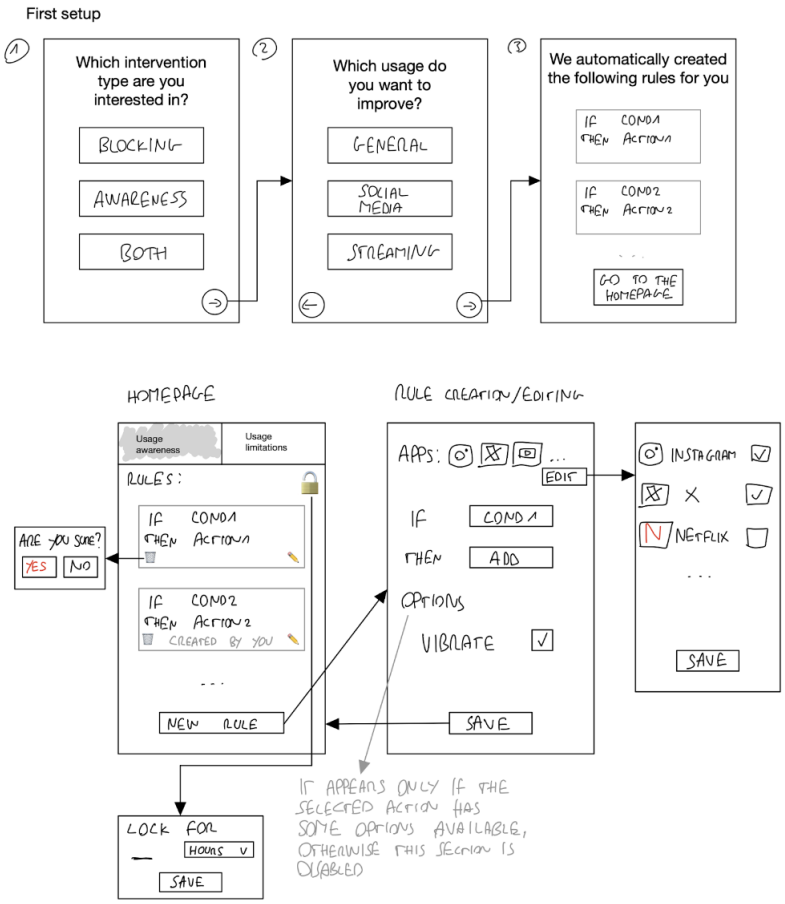


Figure 9: The storyboard of the application.

in (always to consider reactance against one of the two categories) and the usage they want to improve. Based on that, the application automatically creates some rules. The automatic rules will change and adapt based on user behavior during the following weeks.

The homepage remained the same as the first design proposal, just with the addition of a delete button for the rules. Instead, we added some elements to the rule creation/editing page.

Between the proposals and the storyboard, we mapped the desired interventions proposed by the users into triggers and actions, as seen in Table II. Thanks to it, we realized the original interface missed one key element: the applications for which the rule is active. They are, in fact, an implicit and always present trigger, which we overlooked at the beginning.

Furthermore, we realized some actions needed options for more advanced behaviors. For example, some users expressed the need for a timer they cannot bypass, while others asked for a limited number of bypasses (also called tokens). Instead of creating an action for each variant, which would have overwhelmed and confused the user, we decided to merge them into a single action (block the application) they can further customize thanks to options. This section appears only for the actions that need it.

After one last brainstorming session, we changed the final interface prototype. The automatic rule configuration during the setup was not ideal for two reasons. First, it did not allow preferences customization afterward, nor did it allow users not to use them. As we saw in 5.5, it is fundamental for the user to be in control of the tool at any time. Not being able to change ideas about automatic rules could have caused friction in them, potentially leading to the abandonment of the tool. Second, we did not provide any onboarding for the user regarding the rules. During the first use, they may be lost in the interface, not knowing what we mean by awareness and blocks or how to create rules. So, as can be seen in Figure 10, the setup is now used to guide the user toward building their first rule, explaining the procedure step-by-step.

For what concerns the application homepage (Figure 11), we adapted the design to the standard Android design guidelines (using a bottom navigation bar instead of a top bar). Fur-

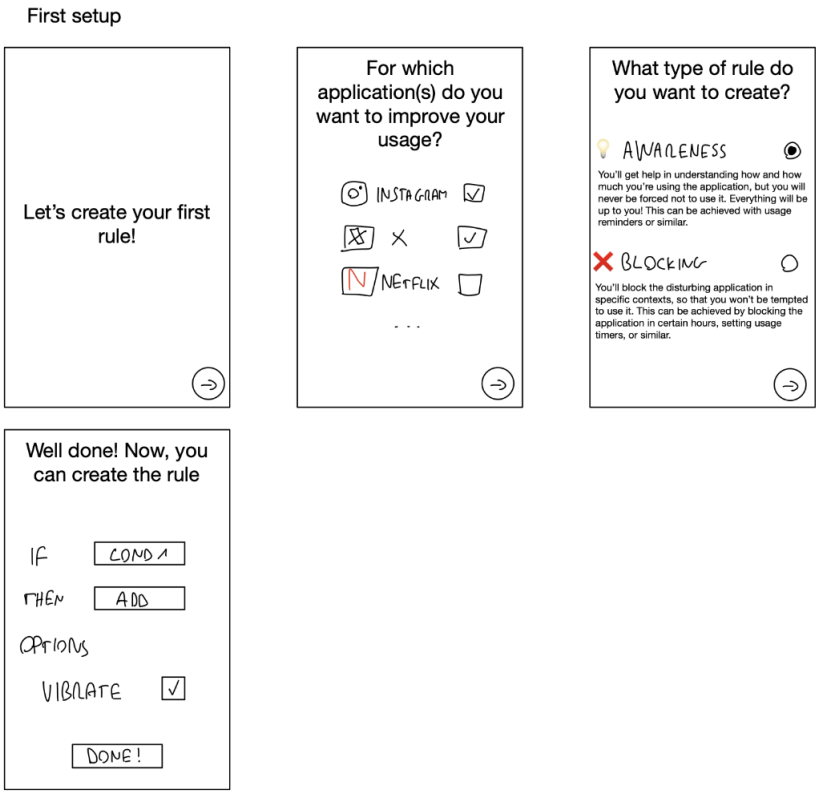


Figure 10: The final setup prototype of the application.

thermore, we added a new section, settings, which contains all the options related to automatic rules. In this way, the user can view and edit them at any time and decide whether to use them.

This design was the final storyboard. As expected, we found ways to improve it while developing, and in some cases, we needed to change our original ideas due to emerging constraints. We will describe them better in 6.5.

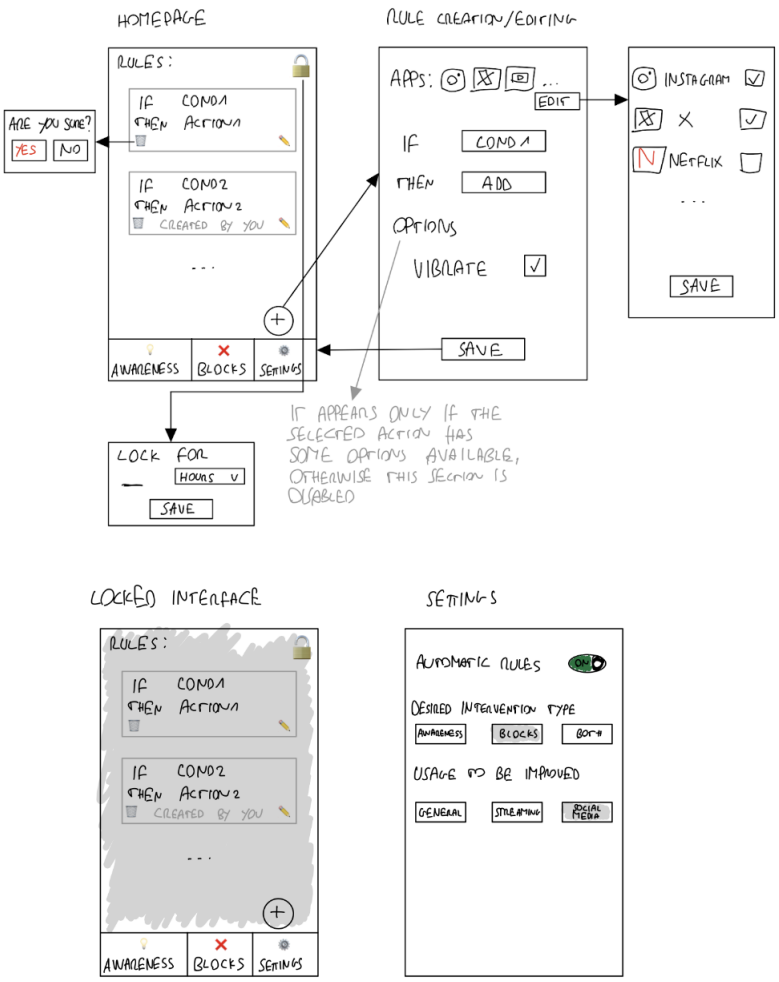


Figure 11: The final homepage prototype of the application.

6.3 Technological Stack

To develop the application, we used Jetpack Compose [44]. This toolkit introduced by Google has recently become Android's recommendation for building native UIs, which is why we chose it.

To store the data, we chose the Room Database for multiple reasons. We needed a solution to store structured data (i.e., rules) without affecting performance and privacy. We considered, but discarded, Firestore for the previous reasons. Performing HTTP requests is much slower than retrieving local data and would not make the application work offline. Furthermore, users' data would leave the device, thus opening up privacy concerns related to how the developer uses it and possible data leaks. Room was a safer and more performing option, so we chose it. To store small data, such as the automatic rule preferences, we used a DataStore.

We used Jetpack Compose with a Model-View-ViewModel architecture, again the suggested way to design Android applications. The layout of the whole application (apart from the setup) is a Scaffold, customized adequately for each page and containing a NavHost. Based on the current route, we display the corresponding View. All the views that interact with the data have a ViewModel, which provides the data to the View and updates the Model. Since the View and the Model have different data representations, it also takes care of converting them into the proper format.

We used a service for the backend, i.e., the rule enforcement. The possibilities we considered for it were a Foreground, Background, or Accessibility service (we discarded Bound services because they were not what we were looking for). Although it would have probably been more

straightforward, we did not use an Accessibility service again for privacy reasons. The user may be reluctant to grant accessibility permissions to the application and may also be discouraged by the alarming popup message shown by Android. This permission provides, in fact, much power to the application, and we should grant it only if we fully trust the developer. Considering the delicate data we deal with, we decided not to use it. Ultimately, we went for the Foreground services because they keep running even if the application is killed (which is not the case for Background ones), thus allowing us to enforce the rules without asking the user to open the application continuously. Apart from the possible reactance, it would also have been a way of bypassing the interventions.

The service starts a Thread, which runs once per second. Each time, it retrieves the application in the foreground by interacting with UsageStatsManager [45]. It then checks if there are any active rules for it. If so, it enforces the rule action. The following pseudocode provides a high-level view of the previously described behavior.

```
val app = getForegroundApp()
rules.ruleList.forEach {
    if(
        it.applications.contains(currentApp) &&
        isRuleActive(it, currentApp)
    ){
        performAction(it, currentApp)
    }
}
```

The service starts when the smartphone boots or the application launches. In this way, the rule enforcement is active in any case, even if the user does not open the application.

6.4 Triggers and Actions

We extracted the list of triggers and actions to include in our application from the participants' desired interventions (5.4). Given the high number of interventions proposed, they cover many situations. The triggers we included are the following.

- Usage time of an application.
- Time of the day.
- Application launched.
- Application currently opened.
- Application closed.
- N minutes scrolling passed.
- Smartphone unlocked.
- Event on the calendar.

The proposed triggers do not belong exclusively to one of the two categories (awareness or blocks), so they are the same in both cases. For the actions, we extracted the following list instead.

- Block the application. (Blocks)
- Show a message to feel guilty. (Awareness)

- Show usage reminders. (Awareness)
- Delay the swipe gesture. (Blocks)
- Insert pauses. (Blocks)
- Blur/filter out content with a specific topic. (Blocks)
- Disable scrolling. (Blocks)
- Popup message to prompt the user to their goal for the usage session. (Awareness)
- Worsen the experience. (Blocks)
- Popup message to prompt the user to the desired usage time. (Blocks)
- Increase battery consumption. (Blocks)
- Remove the application from the home screen. (Blocks)
- Make the phone vibrate. (Blocks)
- Turn the screen grayscale. (Blocks)
- Popup message to prompt the user if they really want to use it. (Awareness)
- Disable notifications. (Blocks)
- Turn off the ringtone. (Blocks)

Since the application is a prototype, we implemented only a limited set of triggers and actions. In particular, we checked the desired interventions list and picked the three most popular ones. As can be seen in 5.4, the two most requested interventions are by far "An application timer without the option to ignore or change it during the day" and "Get notifications that

tell you how much you have been using the application or the smartphone that day." For the latter, since rules are related to a set of applications, we implemented only notifications related to the application usage.

For the third one, many interventions were asked once or twice without a clear preference. We eventually decided to implement "An intervention that purposely worsens the user experience." This idea came up twice in two different co-design sessions, which makes it more interesting because the result of those sessions is a selection of the most critical interventions for the users. Thus, it is not just a random idea or something they said lightly but something they filtered among all their proposals.

The first intervention required the implementation of the triggers "Time of the day" and "Usage time of the application" and the action "Block it." The second uses "Usage time of the application" and the action "Show usage reminders." The third uses the same triggers as the first one and "Worsen the experience" as action.

There are, of course, many ways of worsening the usage experience, and some of them were suggested by the participants, like increasing the content loading time or making the smartphone vibrate. After considering the expressed desires and the ease of implementation, we decided this action reduces the brightness to a minimum and makes the smartphone vibrate. All the side effects disappear as soon as the application is closed. Those changes should make the user's life much more challenging.

6.5 Final Application

The final application maintained most of the design choices described in 6.2. However, we needed to make some changes, either because we did not consider some aspects or because we found ways to improve the interface.

The first significant change regards the setup. Before creating the first rule, the user must grant some permissions, which are needed to make the service and the interventions work. In particular, the application requires the following permissions.

- **Show notifications**. We need it to display the usage reminders.
- **Settings modification**. We need it to override the current brightness value.
- **Display over other apps**. We need it to close other applications. We perform it by creating an intent with ACTION_MAIN as action and CATEGORY_HOME as category. By launching it, the homepage is displayed, thus closing the application.
- **Usage access**. We need it to access the users' usage statistics, such as how long they used each application that day.

For each of them, we created a proper page in the setup, explaining why we needed it (to improve the transparency toward the user) and providing a shortcut to the settings page where it is possible to grant it. After this initial phase, the first rule tutorial appears.

The homepage also has some slight differences. The first change is the short explanatory sentence added below the page title. We indeed explain what we mean with awareness and

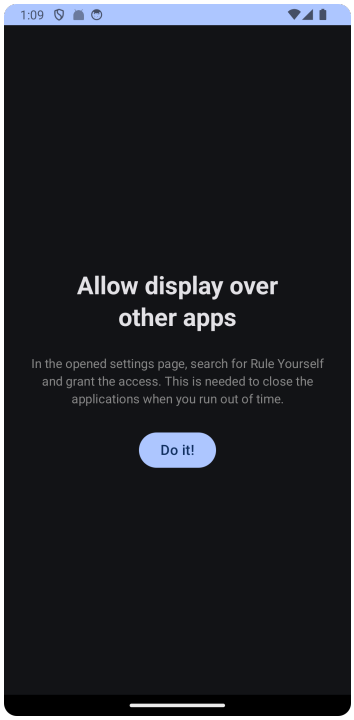


Figure 12: One of the permission request pages.

blocking interventions during the setup, but the users may not read it or forget it after some time. So, we decided to summarize what each section does briefly.

We also made the options visible in the rule card to improve their accessibility. If the rule has some options available, and if the option has a meaningful value (true for boolean, a number above 0 for integers), clicking on the card will display them below the action.

Last, we implemented a lock for each rule instead of a unique global lock. We realized the solution proposed in the storyboard would have been too restrictive, potentially leading to friction. Since the lock cannot be bypassed or edited, it is a strong constraint. The user

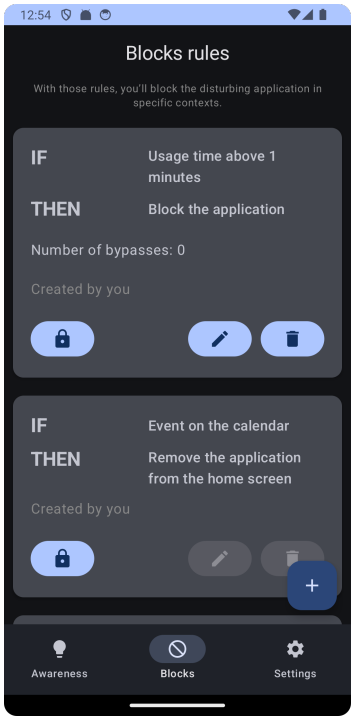


Figure 13: The final app homepage.

could set it by accident or do it intentionally but later realize they want to modify another rule. To limit those issues, we thought of this solution. When a rule is locked, the edit and delete buttons are disabled, thus leaving it in a view-only state.

The last section changed was the one used to create or edit a rule. We already discussed the need for action options, but while implementing the interventions, we realized some triggers need customization, too. For example, the trigger "Time of the day" will require the user to insert the start and end time. Similarly, "Usage time" will need to know if the intervention

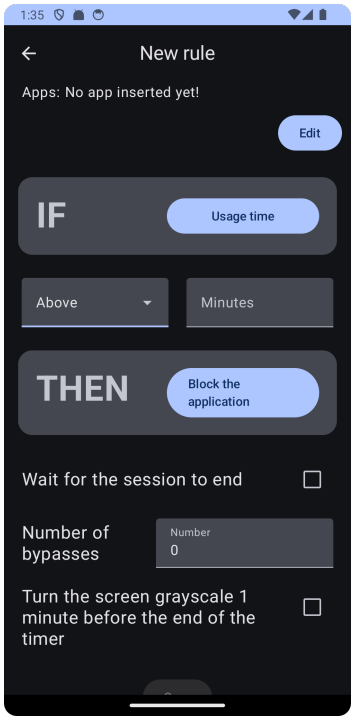


Figure 14: The final rule creation or editing interface, with trigger customization and options.

must be triggered when we reach a threshold (above) or each fixed period (every). To consider this, we added a customization part below the trigger, visible in Figure 14.

CHAPTER 7

LIMITATIONS AND FUTURE WORK

Our work has some limitations. Despite our efforts, we could include only a limited number of participants who were not from Italy (3 out of 39). While their opinions and answers did not radically differ from the ones of the others, involving more people from other countries in the study would have allowed us to generalize the results further.

The most significant limitations, however, come from the application. As explained in the previous chapter, our DSCT was a prototype, with only a subset of the overall functionalities working. We made this choice because our goal was not to implement a fully working product but rather to show in a practical way how the insights of the study and the theoretical framework can lead to a DSCT that can potentially maintain its benefits in the long term. We still added to the interface all the elements that should be present in the final application, like options, so that a future user study would still be feasible.

For what concerns automatic rules, we have already explained their usefulness in the previous chapters. Implementing a fully working version of them would have been difficult, considering the complexity of the task and the fine-tuning required to avoid causing reactance or having no effect on the users. We thought about mocking this feature to give them the perception of how it should work, but, in the end, we decided not to. Users do not have control over automatic mechanisms; if they do not work well, they annoy them rather than improve their experience.

The risk was having all the users disable them for this reason, so we decided to leave it in the interface without any effect. This choice is another limitation.

Implementing all the triggers and actions is endless work. Our prototype included only the triggers and actions extracted from our study. Although they already cover many scenarios, there could be other triggers and actions as needed. For this reason, it could be worth understanding which ones are most in demand before proceeding with the implementation. This idea was another reason behind our choice of creating a prototype instead of a fully working product. Future studies could research the most needed triggers and actions for digital wellbeing.

Another future study could evaluate the effectiveness of our prototype. Our design choices were supported by our research and grounded by the HAM. Still, it could be worth analyzing whether they effectively improve the long-term effectiveness of the tool and if users would prefer something different.

CHAPTER 8

CONCLUSION

This thesis aims to contribute to the delicate yet critical field of digital wellbeing. Initially, we provided an overview of why this domain has become so relevant in recent years and its most significant milestones. We then looked at the HCI efforts for DSCTs and EUD to give a clearer picture of what researchers are trying to improve and the proposed solutions.

We ran a user study involving 39 people with different backgrounds and age groups to collect their experiences and needs to improve their digital wellbeing. The study comprised focus groups and co-design sessions, providing meaningful insights.

We found a theoretical framework, the Habit Alteration Model, which describes how DBCIs can improve user behavior in the long term by changing their habits. Both the theories behind it and the model itself allowed us to understand what users genuinely mean and need with their desired interventions.

Inspired by the theory and the study results, we designed and developed an Android application that allows users to implement rule-based interventions. This solution, which implements Trigger-Action Programming, enables full customization, thus reducing user friction to a minimum and strengthening their commitment toward behavioral change intentions. Those two factors solve two primary reasons behind DSCT abandonment, improving their long-term effectiveness.

Although the efficacy of our solution is yet to be tested, we hope our work contributes significantly toward a paradigm shift in the digital wellbeing field, i.e., the introduction of EUD in DSCTs. The previously described benefits would make the user protagonist in their behavior change, designing themselves, with technological aid, their interventions. Only the user can know what they need, and we should support them rather than provide a premade solution.

Appendix

ORIGINAL MAIN QUESTIONS

1. What are the usage aspects of your phone that you'd like to improve? Why?
2. Tell me about a moment last week when you regretted a phone usage session.
3. Tell me about a moment in the previous week when you used an "addictive" application meaningfully.
4. Tell me about your experience with DSCTs. How long did you use them?
5. Why did you stop using a certain DSCT? (if someone stopped using it)
6. What are the aspects that, in your experience, affect the long-term efficacy of DSCT?
7. Do the available statistics allow you to be fully aware of your phone usage? Would you prefer something different?
8. Do you think the common DSCT implementations (timers, lockout mechanisms, etc.) fully help you use addictive applications less/in a meaningful way? If not, what's missing?
9. Should your current context (at home, outside, etc.) be considered by DSCT? If so, how?
10. Have you ever felt like the DSCT wasn't supporting you in some moments and rather was an obstacle? Why?
11. Tell me about your most recent experience with rule creation/personalization tools (if any).

Appendix

12. Are rule personalization tools understandable? If not, why? How could they improve ease of use and clearness?
13. Tell me about when you had difficulties using rule personalization tools. (if any) How could tools that include personalization rules better support usability?
14. Can you enforce the desired intervention and behaviors with the available DSCTs? Can you make examples of what works for you and what doesn't?
15. If you could design on your own an ideal DSCT, what should it allow you to do?
16. Which DSCT aspects would better support you in your everyday life?
17. Which DSCT aspects would make it likely to use it in the long term?

CITED LITERATURE

1. Google Trends for Digital Wellbeing. https://trends.google.com/trends/explore?date=all&q=/g/11jjl838t0&hl=en. Accessed: 2024-05-02.

2. Forest. https://forestapp.cc/. Accessed: 2024-06-16.

3. Google and the Rise of 'Digital Well-Being'. https://www.wired.com/story/google-and-the-rise-of-digital-wellbeing/. Accessed: 2024-05-01.

4. iOS 12 introduces new features to reduce interruptions and manage Screen Time. https://www.apple.com/newsroom/2018/06/ios-12-introduces-new-features-to-reduce-interruptions-and-manage-screen-time/. Accessed: 2024-05-01.

5. Roffarello, A. M. and De Russis, L.: Achieving digital wellbeing through digital self-control tools: A systematic review and meta-analysis. ACM Trans. Comput.-Hum. Interact., 30(4), sep 2023.

6. Thaler, R. H. and Sunstein, C. R.: Nudge: Improving Decisions About Health, Wealth, and Happiness. New Haven, CT, Yale University Press, 2008.

7. Monge Roffarello, A. and De Russis, L.: Nudging users towards conscious social media use. In Proceedings of the 25th International Conference on Mobile Human-Computer Interaction, MobileHCI '23 Companion, New York, NY, USA, 2023. Association for Computing Machinery.

8. StayFree. https://stayfreeapps.com/. Accessed: 2024-06-17.

9. AppBlock. https://appblock.app/. Accessed: 2024-06-17.

10. Oneplus' Zen Mode. https://www.oneplus.com/global/blog/product/zen-mode-explained. Accessed: 2024-06-16.

11. Dark Patterns. https://www.deceptive.design/. Accessed: 2024-05-01.

12. Mildner, T., Savino, G.-L., Doyle, P. R., Cowan, B. R., and Malaka, R.: About engaging and governing strategies: A thematic analysis of dark patterns in social net-

CITED LITERATURE (continued)

working services. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, CHI '23, New York, NY, USA, 2023. Association for Computing Machinery.

13. Bongard-Blanchy, K., Rossi, A., Rivas, S., Doublet, S., Koenig, V., and Lenzini, G.: "i am definitely manipulated, even when i am aware of it. it's ridiculous!" - dark patterns from the end-user perspective. In Proceedings of the 2021 ACM Designing Interactive Systems Conference, DIS '21, page 763–776, New York, NY, USA, 2021. Association for Computing Machinery.

14. Wacks, Y. and Weinstein, A. M.: Excessive smartphone use is associated with health problems in adolescents and young adults. Frontiers in Psychiatry, 12, 2021.

15. Lanaj, K., Johnson, R. E., and Barnes, C. M.: Beginning the workday yet already depleted? consequences of late-night smartphone use and sleep. Organizational Behavior and Human Decision Processes, 124(1):11–23, 2014.

16. Lyngs, U., Lukoff, K., Slovak, P., Seymour, W., Webb, H., Jirotka, M., Zhao, J., Van Kleek, M., and Shadbolt, N.: 'i just want to hack myself to not get distracted': Evaluating design interventions for self-control on facebook. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, CHI '20, page 1–15, New York, NY, USA, 2020. Association for Computing Machinery.

17. Krasnova, H., Wenninger, H., Widjaja, T., and Buxmann, P.: Envy on facebook: a hidden threat to users' life satisfaction? In Proceedings of the 11th International Conference on Wirtschaftsinformatik (WI2013), Universität Leipzig, Germany, 2013. Bielefeld University.

18. Verduyn, P., Lee, D. S., Park, J., Shablack, H., Orvell, A., Bayer, J., Ybarra, O., Jonides, J., and Kross, E.: Passive Facebook usage undermines affective well-being: Experimental and longitudinal evidence. Journal of Experimental Psychology: General, 144(2):480–488, 2015.

19. Holte, A., Giesen, D., and Ferraro, F.: Color me calm: Grayscale phone setting reduces anxiety and problematic smartphone use. Current Psychology, 42:6778–6790, 03 2023. Accepted: 18 June 2021, Published: 26 June 2021.

20. Davis, K., Slovak, P., Landesman, R., Pitt, C., Ghajar, A., Schleider, J. L., Kawas, S., Perez Portillo, A. G., and Kuhn, N. S.: Supporting teens' intentional social media use through interaction design: An exploratory proof-of-concept

CITED LITERATURE (continued)

study. In Proceedings of the 22nd Annual ACM Interaction Design and Children Conference, IDC '23, page 322–334, New York, NY, USA, 2023. Association for Computing Machinery.

21. Ko, M., Yang, S., Lee, J., Heizmann, C., Jeong, J., Lee, U., Shin, D., Yatani, K., Song, J., and Chung, K.-M.: Nugu: A group-based intervention app for improving self-regulation of limiting smartphone use. In Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, CSCW '15, page 1235–1245, New York, NY, USA, 2015. Association for Computing Machinery.

22. Schwartz, R., Monge Roffarello, A., De Russis, L., and Apostolellis, P.: Reducing risk in digital self-control tools: Design patterns and prototype. In Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, CHI EA '21, New York, NY, USA, 2021. Association for Computing Machinery.

23. Lukoff, K., Lyngs, U., and Alberts, L.: Designing to support autonomy and reduce psychological reactance in digital self-control tools, 2022.

24. Lieberman, H., Paternò, F., Klann, M., and Wulf, V.: End-User Development: An Emerging Paradigm, pages 1–8. Dordrecht, Springer Netherlands, 2006.

25. Namoun, A., Daskalopoulou, A., Mehandjiev, N., and Xun, Z.: Exploring mobile end user development: Existing use and design factors. IEEE Transactions on Software Engineering, 42(10):960–976, 2016.

26. Corno, F., De Russis, L., and Monge Roffarello, A.: Empowering end users in debugging trigger-action rules. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI '19, page 1–13, New York, NY, USA, 2019. Association for Computing Machinery.

27. Manca, M., Paternò, F., Santoro, C., and Corcella, L.: Supporting end-user debugging of trigger-action rules for iot applications. International Journal of Human-Computer Studies, 123:56–69, 2019.

28. Brackenbury, W., Deora, A., Ritchey, J., Vallee, J., He, W., Wang, G., Littman, M. L., and Ur, B.: How users interpret bugs in trigger-action programming. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI '19, page 1–12, New York, NY, USA, 2019. Association for Computing Machinery.

CITED LITERATURE (continued)

29. Corno, F., De Russis, L., and Roffarello, A. M.: A high-level approach towards end user development in the iot. In Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, CHI EA '17, page 1546–1552, New York, NY, USA, 2017. Association for Computing Machinery.

30. Corno, F., De Russis, L., and Monge Roffarello, A.: Taprec: supporting the composition of trigger-action rules through dynamic recommendations. In Proceedings of the 25th International Conference on Intelligent User Interfaces, IUI '20, page 579–588, New York, NY, USA, 2020. Association for Computing Machinery.

31. Corno, F., De Russis, L., and Monge Roffarello, A.: Heytap: Bridging the gaps between users' needs and technology in if-then rules via conversation. In Proceedings of the International Conference on Advanced Visual Interfaces, AVI '20, New York, NY, USA, 2020. Association for Computing Machinery.

32. Huang, J. and Cakmak, M.: Supporting mental model accuracy in trigger-action programming. In Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '15, page 215–225, New York, NY, USA, 2015. Association for Computing Machinery.

33. Zhao, V., Zhang, L., Wang, B., Littman, M. L., Lu, S., and Ur, B.: Understanding trigger-action programs through novel visualizations of program differences. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, CHI '21, New York, NY, USA, 2021. Association for Computing Machinery.

34. Ur, B., Pak Yong Ho, M., Brawner, S., Lee, J., Mennicken, S., Picard, N., Schulze, D., and Littman, M. L.: Trigger-action programming in the wild: An analysis of 200,000 ifttt recipes. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, CHI '16, page 3227–3231, New York, NY, USA, 2016. Association for Computing Machinery.

35. Steen, M., de Boer, J., Kuiper-Hoyng, L., and Visser, F. S.: Co-design: practices, challenges and lessons learned. In Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services, MobileHCI '08, page 561–562, New York, NY, USA, 2008. Association for Computing Machinery.

36. Braun, V. and Clarke, V.: One size fits all? what counts as quality practice in (reflexive) thematic analysis? Qualitative Research in Psychology, 18(3):328–352, 2021.

CITED LITERATURE (continued)

37. Pinder, C., Vermeulen, J., Cowan, B. R., and Beale, R.: Digital behaviour change interventions to break and form habits. ACM Trans. Comput.-Hum. Interact., 25(3), jun 2018.

38. Orbell, S. and Verplanken, B.: The automatic component of habit in health behavior: Habit as cue-contingent automaticity. Health Psychology, 29(4):374–383, 2010.

39. Gardner, B.: A review and analysis of the use of 'habit' in understanding, predicting and influencing health-related behaviour. Health Psychology Review, 9(3):277–295, 2015. PMID: 25207647.

40. Neal, D. T., Wood, W., and Drolet, A.: How do people adhere to goals when willpower is low? the profits (and pitfalls) of strong habits. Journal of Personality and Social Psychology, 104(6):959–975, 2013.

41. de Wit, S. and Dickinson, A.: Associative theories of goal-directed behaviour: a case for animal–human translational models. Psychological Research, 73:463–476, 2009.

42. Muraven, M. and Baumeister, R. F.: Self-regulation and depletion of limited resources: Does self-control resemble a muscle? Psychological Bulletin, 126(2):247–259, 2000.

43. Mobile Operating System market share worldwide. https://gs.statcounter.com/os-market-share/mobile/worldwide. Accessed: 2024-05-07.

44. Jetpack Compose. https://developer.android.com/develop/ui/compose. Accessed: 2024-05-08.

45. UsageStatsManager documentation. https://developer.android.com/reference/kotlin/android/app/usage/UsageStatsManager. Accessed: 2024-05-08.

46. MacLeod, C., Koster, E. H. W., and Fox, E.: Whither cognitive bias modification research? commentary on the special section articles. Journal of Abnormal Psychology, 118(1):89–99, 2009.

47. Alexa Routines. https://www.amazon.com/alexa-routines/b?node=21442922011. Accessed: 2024-06-16.

48. Ifttt. https://ifttt.com/. Accessed: 2024-06-16.

VITA

NAME	Mattia Chiarle
EDUCATION	
	Master of Science in "Computer Science", University of Illinois at Chicago, Sep 2022, USA
	Master of Science in "Computer Engineering", Sep 2022, Polytechnic of Turin, Italy
	Bachelor's Degree in Computer Engineering, Jul 2022,
	Polytechnic of Turin, Italy
LANGUAGE SKILLS	
Italian	Native speaker
English	Full working proficiency
	2022 - IELTS examination (8.0/9)
	A.Y. 2023/24 One Year of study abroad in Chicago, Illinois
	A.Y. 2022/23. Lessons and exams attended exclusively in English
SCHOLARSHIPS	
Spring 2023	Italian scholarship for TOP-UIC students