TU WIEN Informatics

Watchful Detox

Smartwatches und Digital Detox?
Eine Feldstudie

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Media and Human-Centered Computing

eingereicht von

Fabian Pechstein, BSc
Matrikelnummer 0726104

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Univ.Prof.Dr. Florian Michahelles
Mitwirkung: Univ.Ass.in Ambika Shahu, Msc

Wien, 17. Juli 2024

Fabian Pechstein	Florian Michahelles

Informatics

Watchful Detox

Smartwatches for Digital Detox Activities? A Field Deployment Study

DIPLOMA THESIS

submitted in partial fulfillment of the requirements for the degree of

Diplom-Ingenieur

in

Media and Human-Centered Computing

by

Fabian Pechstein, BSc
Registration Number 0726104

to the Faculty of Informatics

at the TU Wien

Advisor: Univ.Prof.Dr. Florian Michahelles
Assistance: Univ.Ass.in Ambika Shahu, Msc

Vienna, July 17, 2024

Fabian Pechstein   Florian Michahelles

Erklärung zur Verfassung der Arbeit

Fabian Pechstein, BSc

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.
Ich erkläre weiters, dass ich mich generativer KI-Tools lediglich als Hilfsmittel bedient habe und in der vorliegenden Arbeit mein gestalterischer Einfluss überwiegt. Im Anhang „Übersicht verwendeter Hilfsmittel“ habe ich alle generativen KI-Tools gelistet, die verwendet wurden, und angegeben, wo und wie sie verwendet wurden. Für Textpassagen, die ohne substantielle Änderungen übernommen wurden, haben ich jeweils die von mir formulierten Eingaben (Prompts) und die verwendete IT- Anwendung mit ihrem Produktnamen und Versionsnummer/Datum angegeben.

Wien, 17. Juli 2024

Fabian Pechstein

Danksagung

Ambika und Florian
Für die Unterstützung und Ratschläge.

Alex
Für deine Geduld mit meinen Projekten. Ich kann mir keine bessere Partnerin vorstellen.

Meinen Eltern
Für die Kuchen und das beständige Fragen, ob ich denn fertig mit dem Studium sei.

Den Studienteilnehmern
Herzlichen Dank für das zahlreiche Feedback!

Tropensturm Hermine
Für das produktive Programmieren am Strand.

Zwei Bänderrissen
Für mehr Fokus.

Berti (der Kater)
Für deine Tisch-Gesellschaft und kreative Eingaben.

Acknowledgements

Ambika and Florian
For the guidance and support.

Alex
For your patience with me. I couldn't imagine a better partner.

My parents
For the cakes and everything.

All study participants
Thanks for your time and all the feedback!

Tropical Storm Hermine
For the highly productive coding days on Tenerife.

Torn ligaments
Supination trauma (left and right foot) for more focus.

Berti (the cat)
For your desk companionship and creative input.

Kurzfassung

Mit dem Aufkommen des Internets und des Smartphones war der beinahe sofortige Zugang zu Informationen, Kommunikation und sozialer Interaktion noch nie jemals zuvor so einfach. Gleichzeiting hat sich die tägliche Bildschirmzeit von Mobiltelefonen auf Rekordniveau erhöht und wurde mit Problemen im Bereich der mentalen, physischen und sozialen Gesundheit in Verbindung gebracht. Interventionen, die im Volksmund auch als "Digital Detox" bezeichnet werden, sind beliebt, um dem Einfluss der digitalen Technologie auf die Nutzer entgegenzuwirken. In den letzten Jahren sind digitale Begleiter in Form von Smartphone-Anwendungen, die den Digital Detox Prozess unterstützen, in der Literatur und auf dem Markt weit verbreitet. Im Gegensatz dazu hat der Einsatz von Wearable Technologie wie Smartwatches in dieser Hinsicht bisher nicht viel Interesse geweckt.

Um diese Lücke in der Literatur zu schließen, stellt diese Arbeit eine Prototypimplementierung eines digitalen Detox-Begleiters auf Basis von Referenzdesigns innerhalb eines Smartwatch- und Smartphone-Ökosystems vor.

Der Prototyp wurde in einer kleinen Benutzerstudie ($n = 6$) über zwei Wochen auf seine Fähigkeiten hin evaluiert, die Telefonbildschirmzeit und ausgewählten Anwendungen zu reduzieren. Gleichzeitig wurde die Telefonnutzung der Teilnehmer in Verbindung mit den physiologischen Daten, die von Smartwatch-Sensoren bereitgestellt wurden, beobachtet.

Vielversprechende Ergebnisse zeigen eine kurzfristige Reduzierung in der ersten Woche, allerdings werden die Reduzierungen jedoch im Laufe der Zeit, trotz zusätzlicher Interventions-Mechaniken während der zweiten Woche des Tests, kleiner. Das beobachtete Verhalten beinhaltet Ausweich-Verhalten und verfrühten Abbruch aufgrund der wahrgenommenen Enge des angewendeten Interventionsdesigns. Dennoch erhielt der Smartwatch-Prototyp insgesamt positives Feedback von den Teilnehmern.

Der präsentierte Methode stellt einen Proof of Concept dar, wobei das implementierte Interventionsdesign Mängel sowie technische Probleme aufzeigt und weiter verbessert werden könnte um den Bedürfnissen der Teilnehmer gerecht zu werden und sich besser zu deren Lebensstil zu fügen.

Abstract

With the rise of the internet and smartphones, almost instant access to information, communication, and social interaction never has been easier. However, in the same manner, daily phone screen time has increased to all-time heights and has been linked to issues with mental, physical and social health. Interventions commonly referred to as "digital detox" are popular to counteract the influence which digital technology has over users. In recent years, digital companions in the form of smartphone applications to aid with the detoxification process have been prevalent in the literature and available products. In contrast, the use of wearable technology, such as smartwatches, in this regard has not seen much interest yet.

To address this gap in the literature, this thesis introduces a prototype implementation of a digital detox companion based on reference designs within a smartwatch and smartphone ecosystems.

The prototype has been evaluated in a small user study ($n = 6$) over two weeks on its capabilities to help reduce overall phone and selected application screen time and to observe participants' phone usage alongside physiological data provided by the smartwatch. Promising results show short-term reduction during the first week; the reductions get smaller over time, despite additional interventions employed during the second week of the trial. The observed behaviour involves evasion and abandonment due to the perceived forcefulness of the applied intervention design Yet, the smartwatch prototype received positive overall feedback from participants, thanks to its simplicity and unobstructed nature.

The presented prototype represents a proof of concept, however, the implemented intervention design has shown shortcomings alongside technical issues and should be further improved to meet participants' needs and better fit their lifestyles.

Contents

Kurzfassung **xi**

Abstract **xiii**

Contents **xv**

1 Introduction **1**
1.1 Is a gilded era upon us? 1
1.2 Undergoing digital detoxification 3
1.3 Presenting research questions 4
1.4 Presenting thesis structure 6

2 Review of Literature **9**
2.1 Effect of digital overload 9
2.2 Digital habit formation 11
2.3 Nudges and challenges of self-control tools 13
2.4 An abbreviated history of digital well-being applications used for screen-time reduction 15
2.5 Towards tangible devices and *Wearables* 23
2.6 Regarding Health–monitoring and, activity and condition recognition with *Wearables* 25
2.7 Combining wearable and phone usage data 27
2.8 Guidelines for designing digital detox applications 28
2.9 Going forward 30

3 Methods **33**
3.1 Preliminary interviews 33
3.2 Intervention Design 34
3.3 User interface design considerations 36
3.4 Platform selection for phone and watch 37
3.5 *Watchful* phone application 38
3.6 *Watchful* watch application 42
3.7 *Watchful*'s components communication 44
3.8 Application distribution and installation instructions 44

3.9 User study 46

4 Results **51**
4.1 Impact of the intervention on phone usage 51
4.2 Observations on health data 60
4.3 Observations from survey and interview data 65

5 Discussion **73**
5.1 Factors supporting a reduction of screen time 73
5.2 Factors mitigating a reduction in screen time 75
5.3 Summary and recommendations 79

6 Conclusion **83**
6.1 Summary 83
6.2 Smartwatches potential 84
6.3 Limitations 85
6.4 Future outlook 86
6.5 Closing words 88

List of Figures **91**

List of Tables **95**

Bibliography **97**

A Survey questions **109**

B Interview guide **113**

C Participant data **115**

D Daily usage of “problematic” applications in respect to defined limits **123**

E Timeline analysis **127**

CHAPTER 1

Introduction

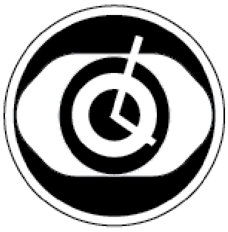


Figure 1.1: *Watchful* application logo: A watchful watcher is ever watching on the watch.

The century-long developments of clocks, their minimisation and affordability allowed to revolutionise the cultural concept of time-keeping and push concepts of efficiency, performance and productivity during the industrial revolution [Lan84, Mar02].

> The watch was an ever-present reminder of the passing of time. [Mar02]

The same motion that helped society to change its pace thanks to an increase in synchronisation did encourage workers to develop a dichotomy of “company time” versus “my time” as described by Landes, a professor of economics [Lan84].

1.1 Is a gilded era upon us?

The age of information has brought humanity never-seen-before benefits. Advances in computing and digitisation brought opportunities to progress even further. With computers that fit into your pocket and are connected to the Internet, access to information has become almost instant. Communication between persons and entities has never been easier, and social media lets us easily create, share and consume content.

However, not all that glitter is gold. For example, in 1976, one negative effect of digital devices and incorporated activities on human beings was described by Joseph Weizenbaum as he observed the following phenomena with his peers:

> *They exist, at least when so engaged, only through and for the computers. These are computer bums, compulsive programmers. They are an international phenomenon.* [Wei76]

The eighties and nineties brought the internet, connecting computers worldwide. Mobile communication devices and social media followed suit, finally finding the current crowning offspring of all these parent technologies: the smartphone. Back in the 70s, programmers were a small community, whereas now, 50 years later, the impacts of these technologies on all society and all age groups can be observed.

In 2013, Ames [Ame13] looked into the techno-social practices of students at the Stanford campus and the skills of the 'digital natives'. Especially how they navigate and multitask social dynamics and personal desires both offline and online were the researcher's concern. One particular participant voiced some criticisms about the constant connectivity that was available through the iPhone:

> *I don't want to be texting all the time instead of enjoying nature and the people around me. People are going to stop seeing the beauty of the world because they're so wrapped up in their iPhones.* [Ame13]

In light of the current state of things — according to data.ai[1], in 2022, users worldwide have spent an average of 4.8 hours per day on their phone. Furthermore, social media and photo & video apps made up 70% of that time—some might say the prediction was not that far off. We see an additional dichotomy of Landes' "my time " [Lan84] into "consumption time" versus "personal time" (with blurry edges).

With every step further into the digital age, we have seen more and more people interacting and using available technologies. In some cases, to a problematic extent, that can threaten their mental and physical health and social sanity and relationships [SH16, KPK+19, YFLL20, BM21, CBK10, MHZ+19].

With the rise of perceived mental load and stress through ubiquitous information availability and connectivity, digital detox activities have become widely popular. The definition provided by the Oxford Dictionary[2] refers to:

> *a period of time when a person does not use digital devices such as smartphones or computers, especially to reduce stress and relax*

[1] data.ai's report on state-of-mobile 2022, https://www.data.ai/en/insights/market-data/state-of-mobile-2022–Accessed August 2023

[2] Oxford Dictionary Digital on "digital detox", https://www.oxfordlearnersdictionaries.com/definition/english/digital-detox–Accessed July 2023

This definition focuses primarily on devices, although it can be extended generously to other digital services, such as social media or entertainment platforms.

1.2 Undergoing digital detoxification

"Digital detox" can also be understood as regaining control of "consumption time" and re-increasing the time available for truly personal use.

The methods of achieving digital detox are manifold. They can range from recommendations or rules to sophisticated technical solutions that help with monitoring, usage and app restrictions or offer gamification incentives [LLS+19, MRDR19] Popular choices are smartphone applications, such as *Digital Wellbeing* (comes with Android), *Screen Time* (comes with iOS), or one of the many options available in Google's Play or Apple's App Store.

Digital detox applications usually employ various features to assist struggling users. Lyngs et al. have identified four feature clusters as part of digital detox applications to introduce behaviour change: Blocking/removal, self-tracking, goal-advancements and reward/punishments [LLS+19].

There are arguably specific issues with habits: mainly, it takes time and effort to adjust habits since changing behaviour is a complex process [MRJ+13]. Achieving short-term results with these apps might be easy. However, the tricky part is building sustainable healthy habits. What might work for one individual might not work for the next one. Choosing the right detox approach and application with the features for your personal needs has become truthfully difficult with the various offerings from all the available app stores.

Yet the question arises: Why rely on the very thing to monitor you that you try to get away from in the first place? Some might call it paradoxical.

Smartwatches have been termed the next killer product after the smartphone [JKC16, RPP14, SHN+17] with a market volume of 134 million units in 2023 and predicted growth of almost 30% over the next five years[3].

Smartwatches can change the way users interact with their phones and might reduce phone dependency [CCB15, VSVB+17]. Since the smartwatch is an intermediary between the user and phone, interactions can be designed to require less attention from the user to help with micro-tasks or provide easy access to information.

However, to date, very few examples of applications of wearables or smartwatches and their usability in digital detox are known. Both in the available literature discussing digital detoxification, nudging, and commercially available products (see Chapter 2).

[3] Smartwatch market statistics, `https://www.mordorintelligence.com/industry-reports/smartwatch-market`–Accessed August 2023

1.3 Presenting research questions

Research Questions at a glance

- **RQ 1:** How might smartwatches aid with digital detox activities and reduce phone screen time?
  - **RQ 1.1:** How do participants react to alarm nudges on the watch, and can they help reduce screen time?
  - **RQ 1.2:** What is the user experience of having their usage data information accessible through a watch face?
  - **RQ 1.3:** What kind of behaviour modification in participants can we observe and measure while using the prototype?
- **RQ 2:** How do usage- and health-data recordings correlate?
- **RQ 3:** Based on these learnings, how might we redesign interventions to reduce screen time using?

Within this thesis I suggest investigating how digital detox interventions could be designed and implemented to assist users in raising usage awareness and reducing screen time with the use of smartwatch technology:

> **RQ 1:** How might smartwatches aid with digital detox activities and reduce phone screen time?

The research approach proposed is to establish an understanding of the current state of research, commonly available and implemented intervention mechanics in similar applications and derive user needs through preliminary user interviews.

Several possible intervention designs are feasible, given modern smartwatches' capabilities and feedback modalities. The most relevant feature of wearing a watch is the easy access to information through its wrist-mounted display or interface, commonly known as the *watch's face*. This unobstructed manner of wearing the device at the wrist allows fast interactions and reading presented details at a glance with low cognitive effort.

To address the current literature gap and help answer this research question, a prototype was developed (see Figure 1.2; this also serves as an example of a watch face). Assuming that, from the starting point of this research, an alarm mechanic to enforce screen-time limits and a watch face will be part of the intervention design proposed in this work, the question can be further divided into the following two sub-questions, which will be evaluated during a user study employing the prototype:

Figure 1.2: *Watchful*'s watch face in action

RQ 1.1: How do participants react to alarm nudges on the watch, and can they help reduce screen time?

RQ 1.2: What is the user experience of having their usage data information accessible through a watch face?

Furthermore, with various data recordings and metrics available that describe participants' behaviour, one additional question arises:

RQ 1.3: What kind of behaviour modification in participants can we observe and measure while using the prototype?

Another aspect of smartwatches is the promising opportunities to modify physical activity behaviour [SAVL+16, GMvO+18, BWO+19, TMM+20] and increase health awareness among their users. In this regard, social media, among others, have additionally been proven valuable as tools for health promotion and education [SSL10, KI13].

However, with tracking data available for physiological information, physical activities and phone usage, another interesting question can be formulated:

RQ 2: How do usage- and health-data recordings correlate?

This leads to the final research question, which tries to summarise all the findings from the prior questions into suggestions or guidelines for future research based on the findings of this thesis:

> **RQ 3:** Based on these learnings, how might we redesign interventions to reduce screen time using?

1.4 Presenting thesis structure

This thesis has been structured in the following way, as visualised in Figure 1.3:

Chapter 2 reviews existing literature using repository academic search engines *Google Scholar*, *IEEE Xplore*, *ACM Digital Library* and *Science Direct* (Elsevier) to identify relevant reference designs and applications as well as the theoretical background (see Chapter 2).

With the identified gap of watch-based interventions in the current body of literature and a review of reference intervention applications to help reduce phone screen time, designing and implementing a prototype was the logical next step. The technical capabilities of mobile operating systems and available smartwatch hardware and software products were evaluated in Chapter 3 to decide on a usable prototype platform (see Section 3.4). Preliminary interviews ($n = 6$) were conducted to get a more grounded understanding of user needs and how others felt about their phone usage, their strategies to manage screen time and the use of smartwatches (see Section 3.1).

Based on these findings and decisions, designs of possible and implementable intervention methods were created or adapted from reference applications (see Section 3.2).

As the next step, an Android phone application was developed with a companion WearOS application to explore and evaluate the proposed intervention designs (see Sections 3.5 and 3.6).

The prototype was evaluated in a pilot study with two participants to gain insights into authentic user experiences and potential technical issues. The revised prototype design was evaluated with a small ($n = 9$) user study to test our proposed intervention scheme and to capture user feedback about the participant's personal experiences during the study (see Section 3.9).

Data from six participants were selected to be further analysed (three participants were excluded due to loss of data or faulty data records). Results are presented in Chapter 4 and further discussed in Chapter 5. Chapter (6) completes the thesis and presents the conclusions and final thoughts.

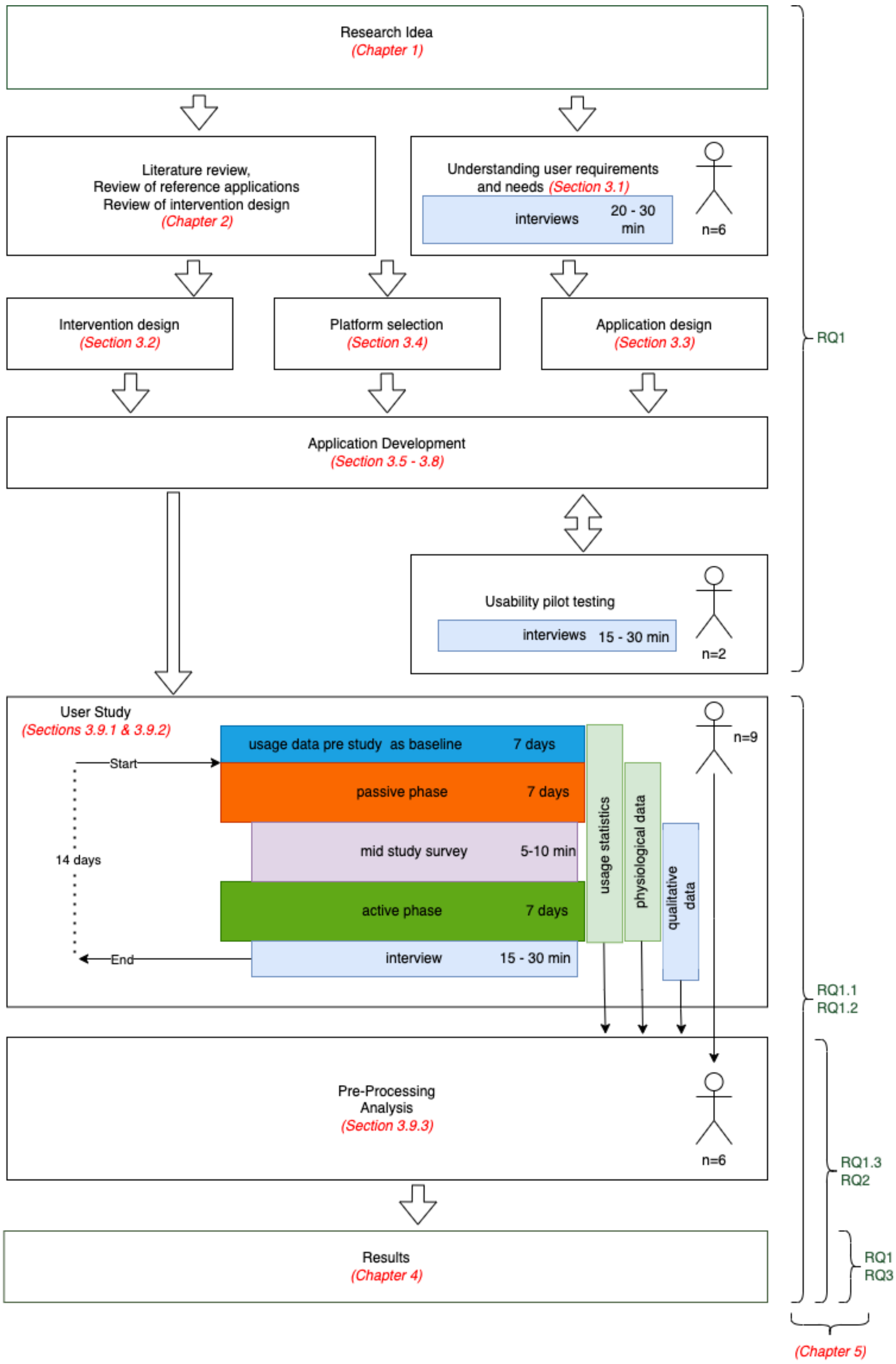


Figure 1.3: Research Design (qualitative data in *blue*, quantitative data in *green* and mixed data in *pink*)

CHAPTER 2

Review of Literature

The following chapter represents a review of the literature discussing the theory of change behaviour, habit-formation, published implementations and designs of smartphone-based applications to help with screen-time reduction.

As reference publishing repositories and search engines, *Google Scholar*, *IEEE Xplore*, *ACM Digital Library*, and *Science Direct* (Elsevier) were used.

2.1 Effect of digital overload

By default, applications like Android's *Digital Wellbeing*[1] or Apple's *Screen Time*[2], to name only the ones that are available from mobile operating system manufacturers, are pre-installed on mobile devices. However, the fact that average daily phone screen time has been increasing over the last few years fosters the assumption that promoting awareness or actively reducing product usage is not a primary interest of smartphone manufacturers, platform providers, application developers or users. Easy access, captivating and frictionless designs, algorithms that find endless relevant content and gamification mechanisms motivate their users to spend more time with applications.

With the rise of digital usage concerns, tools have been provided to monitor, moderate or restrict device usage. In the end, however, it's up to users to build usage awareness and employ them in their personal space.

The underlying root cause of people's extensive preoccupation with digital devices, such as smartphones and digital content ('Internet Addiction' and social media)[KG11], is

[1] Digital Wellbeing, `https://www.android.com/digital-wellbeing/`–Accessed November 2021

[2] Screen Time, `https://support.apple.com/en-gb/guide/iphone/iph24dcd4fb8/ios`–Accessed November 2021

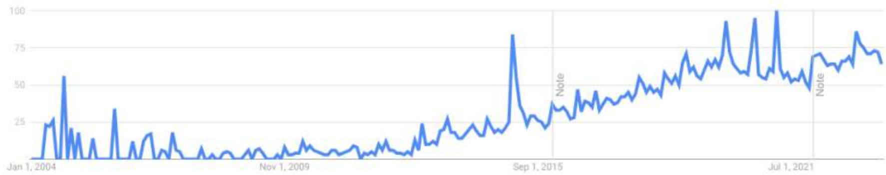


Figure 2.1: Google search trends for the term 'digital detox'

currently still debated [BMLF+15]. Reasons for the predicament might be likewise multifaceted or driven by underlying conditions, for instance, depression or low self-esteem.

Whereas there is a dispute about the origin of addictive behaviour, there exists a kind of consensus regarding therapy: mainly abstinence or moderation of use, among others. This brings us finally to 'digital detox' as a coping strategy to target excessive usage.

In a systematic review by Radtke et al. [RAS+21], the authors try to find evidence of beneficial outcomes with digital detox as an intervention method. Although numerous studies have been considered, general conclusions are diverse and show mixed or contradictory findings regarding detox efficacy. Radtke et al. blame the variety of implementations of digital detox, among other factors, like different measurements of similar outcomes, assessment timings, participant selection, and randomisation of study groups. Further limitations highlighted by the authors are the lack of device-based measurements of users' behaviour that would allow a more reliable assessment than self-reported behaviour and tracking compensation behaviour (i.e. using a non-monitored device to access social media).

Yet, despite this large body of research focusing on the negative side of social media, others found positive effects, e.g. as a tool to get health information or emotional support [SP19, GT20].

An eight-year longitudinal study investigating the relations between social media, depression and anxiety by Coyle et al. found that those participants using social media more often have higher scores in mental health difficulties. But from a within-person point of view, the changes in one (social-media) do not covary in the other (e.g. depression), but instead the authors found no evidence of social media being a risk-factor to developing depression and anxiety in their sample at hand. This lets the authors wonder 'if society is in a moral panic, at least concerning the sheer amount of time spent on social media and mental health' [CRZ+20].

Wadsley and Ihssen suggest that voluntarily restricting social media use for one week shows nuanced and potentially off-setting effects on well-being, as both triggers for negative and positive emotions are removed [WI23].

The worldwide COVID-19 pandemic (starting late 2019) and accompanied by month-long lock-downs during 2020 and 2021 due to quarantine policies increased the speed of digital transformation processes in various sectors (e.g. education [ISVO20] and health [PSS20]).

Sparked new discussions about home office and its impacts on physical [WSRS22] and mental well-being through digital overload [MHF22]. Figure 2.1 shows the worldwide interest in ‘digital detox’ since 2012[3]. As we can see, it’s an upward trend, with the last three spikes (January 2020, September 2020 and March 2021) during the COVID pandemic.

Parry et al. [PlM+23], while investigating the adoption rate of digital detox applications, found that although 80.48% of their participants indicated that they are aware of digital well-being applications, 63.06% of those respondents indicated that they have ceased using (16.12%) or have never (46.94%) used such applications.

2.2 Digital habit formation

A habit is a disposition to perform a given behaviour [NWQ06, Gar15, PVCB18]. Habitual behaviour is the behaviour that results from this impulse [PVCB18]. People are not necessarily unaware of their behaviour but are usually ignorant of the habit’s inner workings, which makes them hard to change [WR16]

Dual system theory is a model to explain how impulses can arise, and decisions are being made. It describes two distinctive cognitive processes: System 1 is the unconscious, fast and parallel pathway. Making use of automatic responses, learned habits and heuristics. System 2 is slower, conscious, and driven by goals, intentions and rules. The latter process is necessary for planning, decision making or overcoming habits and temptations. Hence, to make a new desired behaviour (i.e. healthy) a habit, it needs to make the transformation from a system two process into a system one path [LLS+19, Kah11, PVCB18].

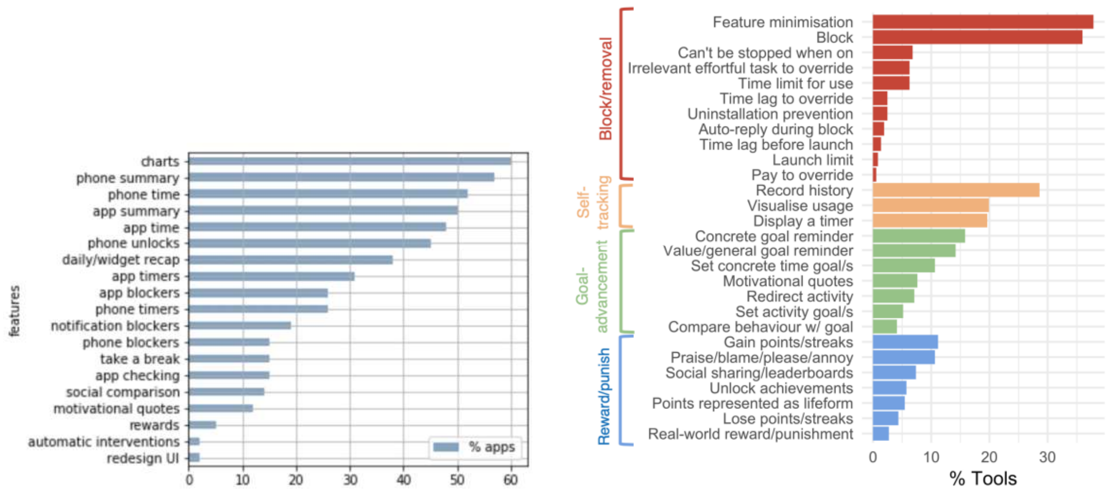
In the context of digital detox, the following recommendations have been made:

Pinder et al. outlined several theoretical approaches to change habits and recommend the implementation of intentions and automation (i.e. context-based) of self-regulation in the habit formation regard [PVCB18].

Monge Rofarello and De Russis [MRDR19] have classified available digital well-being assistants (in Google Play Store only, $n = 42$) in an exploratory study into two categories based on the provided features — self-monitoring and interventions. Interventions can occur on the phone or app level (i.e. timers and blockers) and additional mechanics like rewards or motivational quotes. Meanwhile, self-monitoring features describe tracking and data-presentation functionality. Tracking and data visualisation are among the most popular features in the examined applications (see Figure 2.2a), whereas intervention features are more sparsely distributed. This could be explained due to the different approaches and mechanics each of the examined detox assistants implemented.

Based on the findings of their exploratory studies, the authors built their well-being assistant. They evaluated it in a user study, which revealed that despite self-tracking playing an essential role in a successful behaviour change process, it rarely supports the

[3] Google Trends, https://trends.google.com/trends/explore–Accessed August 2023



(a) Distribution of features in digital well-being assistant applications (in Google Play Store only, $n = 42$) [MRDR19]

(b) Functionality codes of digital self-control tools (Google Play Store, Apple App Store and Chrome Extensions, $n = 367$) [LLS+19]

Figure 2.2: Feature and features clusters commonly used in detox applications [MRDR19, LLS+19]

formation of sustainable behaviour since the removal of the monitoring entity participants mostly relapse into beaten tracks.

Consequently, the authors identify a need to be grounded in habit formation research to provide long-term adherence to healthy behaviour and recommend making use of the contextual-awareness functionality of smartphones to suggest new habits dynamically [MRDR19].

Similar to Pinder et al. [PVCB18], Lyngs et al. [LLS+19] have provided a study of Android and iOS apps and browser extensions that serve as digital well-being and detox assistants. And investigated their overlaps or differences in their implementation of applying dual system theory. Figure 2.2b shows the distribution of features in the examined apps ($n = 367$), which is, to some degree, despite smartphone applications with browser extensions and using different clustering and labels, comparable with the findings of Rofaello and De Russis [MRDR19]. Lyngs et al. identified gaps in available implementations that could be used for habit formation: scaffolding habits instead of blocking mechanisms. Furthermore, they found a lack of the use of delay mechanisms and expectancy components. Which represents how likely the user thinks he will succeed in achieving his goals.

2.3 Nudges and challenges of self-control tools

According to Thaler and Sunstein, a nudge 'is any aspect of choice architecture that alters people's behaviour in a predictable way without forbidding any options or significantly changing their economic incentives' [TS08, p.6]. The idea is to put the healthier option in a relatively nearer location, speaking in abstract terms. For example, providing employees with a free fruit basket next to the snack machine is a nudge.

One intervention method commonly used for behaviour change in digital environments is digital nudges. Weinmann et al. describe the term as 'the use of user-interface elements to guide users' behaviour when they are required to make judgements or decisions in a choice environment [WSB16]. Yet the outcome doesn't just depend on the thoughtful consideration of the available options alone. The design of how the options are presented plays an additional subconscious role. This gives designers or choice architects the capability to influence behaviour following dual-process theory [WSB16, TS08].

Affecting an outcome of a decision can be already done by changing default options (e.g. opt-in v.s. opt-out as the default option) as shown by Johnson and Goldstein [JG03]. Besides settling on defaults, other techniques are available to choice architects — summarised and provided as taxonomy by Münscher et al. [MVS16].

Digital nudges have found their use in multiple applications for digital detox apps, which will be discussed in Section 2.4

Habits represent learned behaviour which can be automatically used by *System 1* processes in the context of dual system theory. Unfortunately, building up habits takes time. *mHealth* user studies implementing interventions to promote more active lifestyles describe three to six months until exercises become habitual behaviour [LVJPW10].

Despite the duration of the intervention activities, the self-imposed strictness and abandonment of self-regulation methods pose another issue when they become unbearable for the user. Schwartz et al. describe four basic design patterns to reduce the overall risk of abandonment [SMRDRA21]. Firstly, they use continuously variable interventions that can be scaled up or down in case of user dissatisfaction or concern. Secondly, bundling of interventions and thus pursuing a mixed strategy of interventions (e.g. summary visualisations and app timers) becomes more likely to succeed. Thirdly, users have been left in control over which intervention they partake in, although they are often disappointed by the results. This can be mitigated by letting an expert prioritise the selection, which might benefit the user's experience and expectancy. Lastly, digital well-being assistants rely on consistent behaviour, which can be challenging due to changes in the technical environment. By implementing these patterns, the authors expect reduced user autonomy, accompanied by a reduced risk of abandonment of the ongoing interventions. A lower risk of intervention discontinuation should support the formation of healthy habits in the long run.

Kovacs et al. [KWB18] additionally propose a rotation of interventions because users tend to ignore them after a while.

Hummel and Maedche have categorised nudges and evaluated their effect sizes in a systematic literature review. They conclude that the nudges might be less effective than anticipated depending on the context (application) and nudge dimension (after Münscher et al. [MVS16]). They found that nudges promoting 'defaults' show significant effect sizes in 62% of the studies. However, in a health-related context, nudges are more likely to use change effort, which shows only half the effect size (median 25%) compared to defaults (median 50%) [HM19].

Ultimately, nudges can always be used and most certainly are used to promote ulterior motives that are not in the best interest (so-called 'dark patterns' in UX design, coined by Brignull[4]). Thaler refers to such malpractices that don't have the user's best interest in mind as 'sludge' and recommends encouraging the public and private sectors to engage in 'sludge cleanup campaigns' [Tha18]. Gold et al. found that transparent behavioural interventions resulted in higher acceptance rates for various fields, except when encouraging physical exercise [GLAO23]. Purohit et al. discussed that although users might not be aware of the psychological mechanisms employed by choice architects [Mil23], they should have the autonomy to follow or unfollow nudges as they please [PBS+23]. Purohit et al. summarised these as a guideline on how nudges are supposed to be [PBS+23]:

1. transparent [GLAO23]
2. easy to opt out. [PH21, PBS+23]
3. designed with the user's well-being in mind [Tha18]

Schwartz [Sch19] provides several criticisms about the nature of self-regulation tools, which partly influence the design patterns mentioned above [SMRDRA21] to limit the risk of intervention abandonment. However, Schwartz proposes that self-control tools must combine the desires of all parties involved in using the tool: users, tool designers, platform owners and developers of blocked applications. The tool represents, consequently, a form of control by others and oneself.

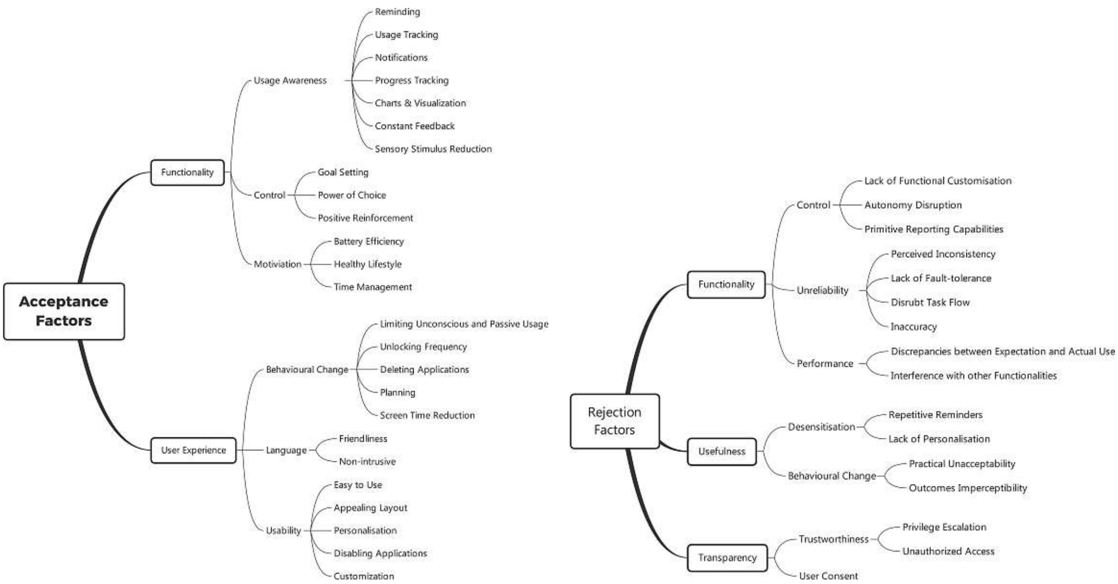
Purohit et al. found that digital nudging adoption is hindered by usability and ethical concerns [PBH20, PH21]. Nudges designed for digital detox often introduce increased efforts to complete a task related to undesired behaviour or add subtle to forceful feedback (e.g. notification) to ongoing activities, thus impacting usability and user experience and eventually leading to disregard or abandonment in the long term.

Ethical and privacy concerns about digital detox tools or nudges analysing usage data, intervening or changing experiences leave a bitter aftertaste of lost control over free choices, surveillance and paternalism [JSD+12, ST03, Wid20].

Almourad et al. evaluated user reviews of *Digital Wellbeing* an *SPACE*[5] and extracted factors influencing user acceptance and rejection towards digital well-being applications

[4]Dark Patters, https://www.deceptive.design/types–Accessed August 2023

[5]SPACE, google play, https://play.google.com/store/apps/details?id=mrigapps.andriod.breakfree.deux–Accessed August 2023



(a) Factors favouring user acceptance (b) Factors leading to user rejection

Figure 2.3: Acceptance/rejections factors according to Almourad et al. [AAS+21]

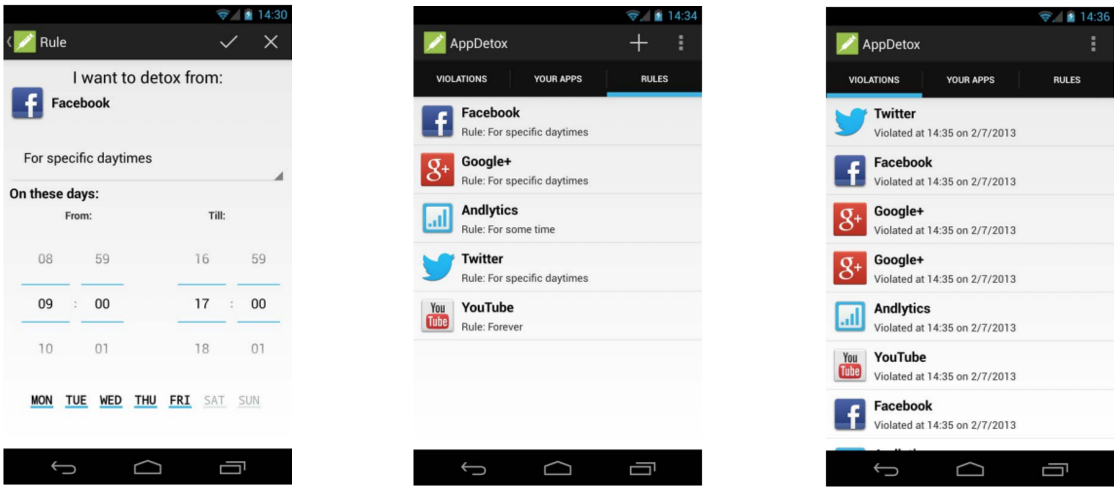
through thematic analysis (see Figures 2.3a and 2.3b). The findings show that implementations of usage awareness positively influence user acceptance. However, the other part of the presented results shows how carefully all parts (user interface, the nudges, language involved, technical aspects) of the application must be designed [AAS+21] to not lead to user rejection.

2.4 An abbreviated history of digital well-being applications used for screen-time reduction

This section overviews the literature's commonly available digital well-being and detox applications and approaches.

The list of applications and approaches is not complete by far, but it serves to showcase examples of users' needs during that period. There is a noticeable switch from just providing tools for monitoring and regulating phone and app usage to investigating the most effective ways to reduce usage and build sustainable habits. In recent years, there appears to be more focus on even more specialised applications using more subtle but visible feedback targeting to help reduce social media or news consumption [PBS+23, OSDE18]

In 2013, Löchtelfeld at al. released an app called **AppDetox** in the Android Play store [LBG13]. The provided features were based on three rules: One would allow users to define time-frames in which opening a specific app was restricted. Another rule would categorically block the starting of apps, and lastly, one would create a countdown during



(a) Rule creation (b) Active rule (c) Rule violations

Figure 2.4: *AppDetox* [LBG13]

which the app can't be used (see Figure 2.4a and 2.4b). Breaking a rule was recorded and displayed in an extra overview (see Figure 2.4c). Creating and maintaining rules was voluntary, and users seemed to have used them in the morning to prepare for the day ahead. The authors analysed the rules and violations of almost 12k users and found the top 3 rules defined were for messaging, social media and browser applications [LBG13].

In 2016, Hiniker et al. proposed the app **MyTime**, which tries to support people in achieving goals related to smartphone non-use [HHKK16]. Three intervention nudges were implemented: Firstly, users were asked what they wanted to achieve today, then they could select from a catalogue of all installed applications apps they wanted to track and set a timer for maximum daily usage. Whenever one of the tracked applications was started, a live notification in the status bar showed a progress bar displaying the time a user had spent with the app regarding the set limit.

If a limit was reached, a UI overlay would appear, displaying the set limit and the defined goal for the day, proposing an alternative to phone use. Users then had the option to dismiss the overlay, to be reminded of the timer again after a short while (5 minutes or 10% of the limit) or to acknowledge the timer's proposal, thus quitting the monitored app and returning to the phone's home screen (see Figure 2.5).

Participants ($n = 33$) reduced daily by an average of 33 minutes (11%) (1 week of baseline and one week of intervention). The timeout feature was partially used to gain information about usage, which was reflected in how users reacted to the 'time's up' notifications. However, participants reduced their use of applications that they considered a 'waste of time'. The feature was not used by all participants equally. Hiniker et al. found that those interested in making contextual changes to phone use engaged more with the aspirations feature, whereas others chose not to engage with this specific feature entirely.

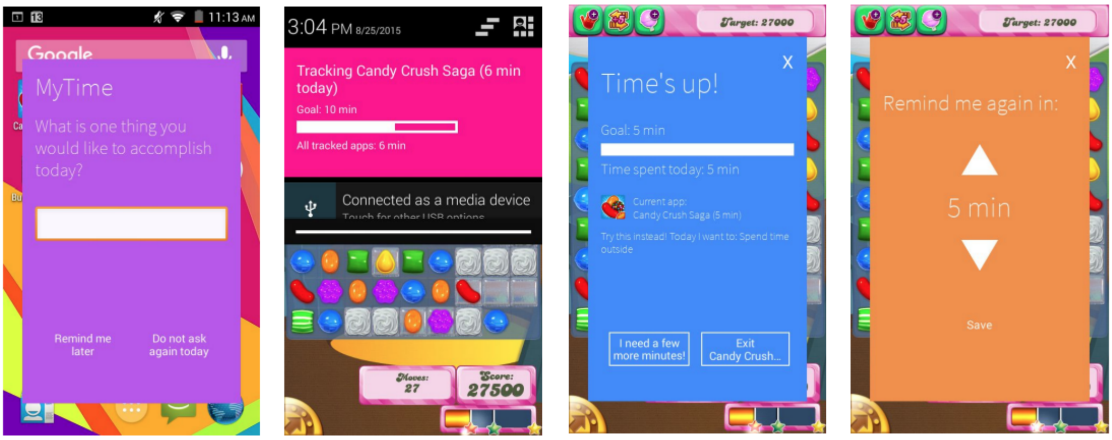
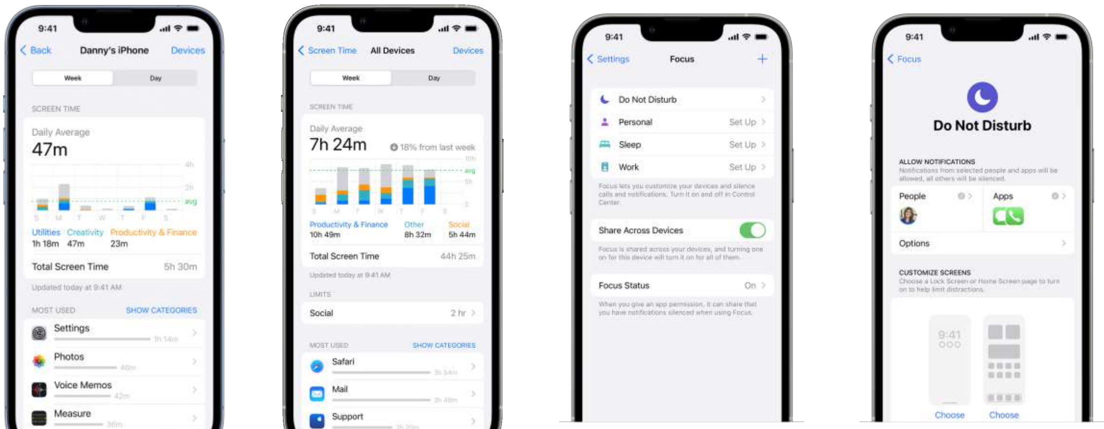


Figure 2.5: *MyTime* by Hiniker et al. [HHKK16]



(a) Overview

(b) Detail overview

(c) Focus mode settings

(d) Bedtimer mode settings

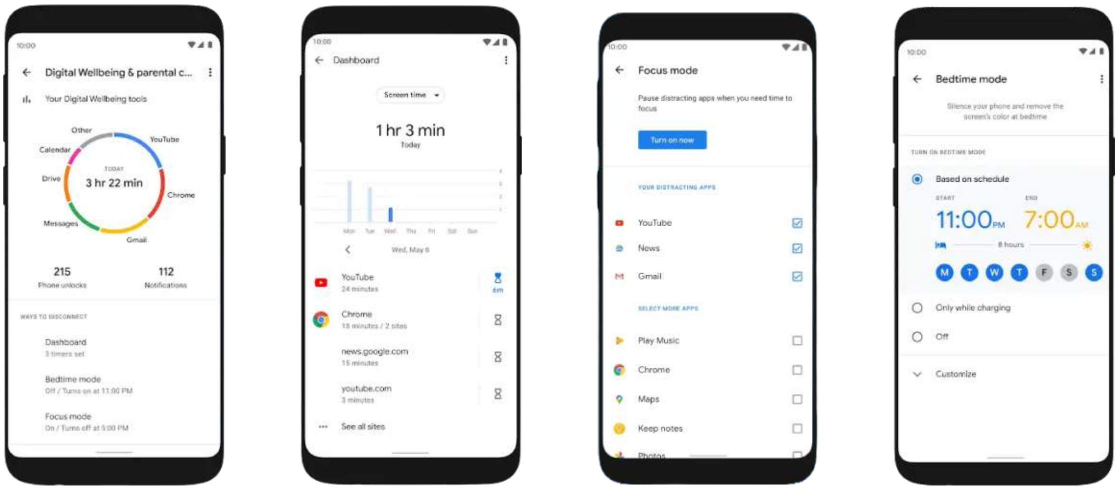
Figure 2.6: iOS's *Screen Time* and *Focus* (iOS 16)

The authors were unsure how the behaviour change fares in the long term [HHKK16].

With iOS 12, released in September 2018, Apple added **Screen Time** to the roster of built-in system features. With this application, it became possible to get detailed usage information, block or limit time spent with certain apps, and even get weekly reports about usage trends in iOS (see Figure 2.6a). Usage data could be gathered over various devices (see Figure 2.6b) within the Apple ecosystem as soon as users activated their Apple ID and opted-in for data collection. iOS 15 (2021) added another application called Focus, which allowed schedules and more advanced rules for do-not-disturb modes for different profiles (personal, sleep, work) (see Figures 2.6c and 2.6d).

Also, starting with 2018, Android got a built-in application called **Digital Wellbeing**[6]

[6] Digital Wellbeing–https://www.android.com/digital-wellbeing/–Accessed June 2023



(a) General view (b) Detail overview (c) Focus control (d) Bedtimer

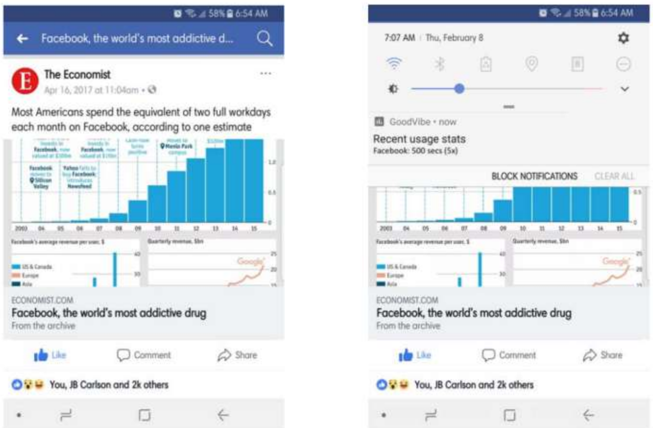
Figure 2.7: Android's *Digital Wellbeing* (Android 13)

responsible for various aspects of digital detox and was initially launched in 2018. The app is deeply integrated into the operating system and other Google applications and provides a public API that developers can use, e.g., to gather usage statistics.

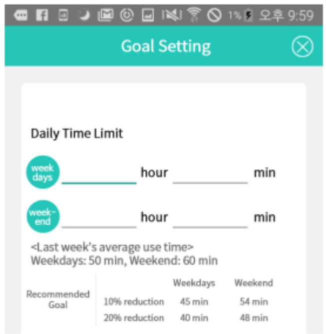
Dashboards and overviews provide information on daily and weekly app usage, phone unlocks and received notifications by app. Historical data was accessible only up to 4 weeks prior, although it's possible to programmatically gather aggregated data for more extended periods. Setting usage time limits for apps and URLs when browsing in Chrome is possible. After exceeding these goals, the app or URL will get blocked and can't be used for the rest of the day unless the limit gets removed. Additionally, a focus mode can be scheduled to pause a user-specified list of applications, disable notifications for the given time frame and display a dialogue when a focus-listed app starts. This leaves users with the choice to discard their try to use the app or use it for a limited time. A so-called bedtime mode can also be scheduled, which silences the phone and all notifications, dims the brightness of the background wallpaper and turns the screen into black-and-white mode.

Given how deep third-party apps need to be integrated into systems to support restricting features, providing default built-in tools in the commonly used mobile operating systems can be seen as a sensible strategy to ensure users' privacy.

Okeke et al. [OSDE18] created an application for Android called **GoodVibe** that provides users with real-time, textual feedback of how much time they have spent on a mobile application. Furthermore, a digital nudge in the form of gentle phone vibrations when a specified daily usage limit for an app is reached (see Figure 2.8a) was added. The application could track and monitor all presently installed applications on the phone. However, the authors focused on Facebook and limited only it's application for the user



(a) *GoodVibe* by Okeke et al. [OSDE18]

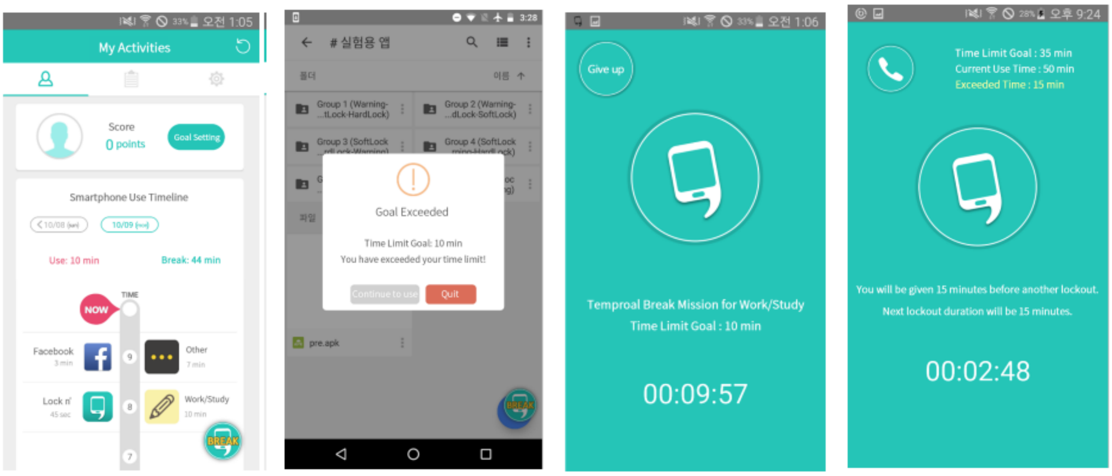


(b) *GoalKeeper*'s alarm configuration screen. Notice the different options for weekdays and weekends [KJKL19].

Figure 2.8: screenshots of *GoodVibe* by Okeke et al. [OSDE18] and *GoalKeeper* by Kim et al. [KJKL19]

study. An MTurk study that split participants ($n = 50$) into three groups (control: no limits, just the feedback ($n = 16$); rigid: fixed limit of 30 minutes and five times opening the app ($n = 29$); personalised: 50% of benchmark's week average usage and start counts ($n = 15$)) was conducted over three weeks. The first week was to capture the baseline, the second would introduce the vibration nudge, and the last week would remove the nudge again to assess the effects of removing the intervention. Okeke et al. found significant reductions between the first and second weeks for the groups using the nudge. However, there have been no significant changes for the last week. Participants reported that the vibrations increased their awareness but were not strong enough to cause them to abandon Facebook entirely, as some participants would continue to use the platform after the vibration had started. A follow-up one year later revealed that the replying participants ($n = 21$) reported feelings of overusing Facebook and simultaneously using no app to help them monitor their use. Even though 90% claimed that it's important to them to regulate time spent on applications.

Kim et al. own application **Goalkeeper** [KJKL19] employed several different mechanisms and proposed comparably more substantial restrictions on their users. Standard features like a timeline (see Figure 2.9a), the capability to define time limits for applications (daily and weekly limits are supported) and utilising the notification bar to provide quick access to usage information were included. The authors, furthermore, introduced a break and lockout mode during which only basic phone functions can be accessed (physical buttons got deactivated). Break mode can be started voluntarily and stopped at any given time (see Figure 2.9c), whereas lockouts are triggered after exceeding defined usage limits for applications (see Figures 2.9d and 2.9b). For the evaluation study, three different levels of severity for the lockout were used: non-locking (no consequences after reaching



(a) One of *GoalKeeper*'s timeline and statistic views

(b) *GoalKeeper*'s interaction lockout dialog providing a choice

(c) *GoalKeeper*'s voluntary break mode to make focusing easier

(d) *GoalKeeper*'s lockout screen after exceeding the limit goal

Figure 2.9: Screenshots of *GoalKeeper* [KJKL19]

limits), weak-locking (increasing longer periods of locking after periods of allowance time) and strong-locking (locking the phone after exceeding limits until midnight). A four-week deployment study was conducted where participants ($n = 46$) were randomly assigned to a lockout-level group after the first baseline week and advised with a 10% and 20% reduction goal based on baseline measurements. Most participants felt pressured by the lockout mechanism and consequently planned their smartphone usage or even diverged to other computer devices. In contrast, participants from the non-locking group naturally ignored the fact that they had reached their limit (some felt disappointed and guilty). Thus, the authors found that the stronger the lockout mechanism, the more the user can reduce their usage time by suppressing the temptation to use their phones. Interestingly, a few participants found exploits that let them leave lockout mode due to software issues, and in turn, their commitment level towards screen time reduction seemed to have weakened. They would rather continue to use their phones. The authors found the short study duration a limitation, as the long-term effects of adoption processes were not observable [KJKL19].

Monge Rofarello & De Russis have implemented **Socialize** (see Figure 2.10) after reviewing 42 available digital detox applications using features that were present in at least 15% of these apps: tracking, data presentation, timers and contextual interventions. For evaluation, 69 participants installed and used the app for three weeks (a one-week baseline and a two-week intervention phase). The authors found a significant reduction in screen time and the number of apps used, in contrast to phone unlocks and the number of app executions. Despite using Socialize, participants constantly checked their phones, but they reduced the time spent on social media significantly. The authors found that their app and others alike can be useful to reduce screen time. However, interventions are

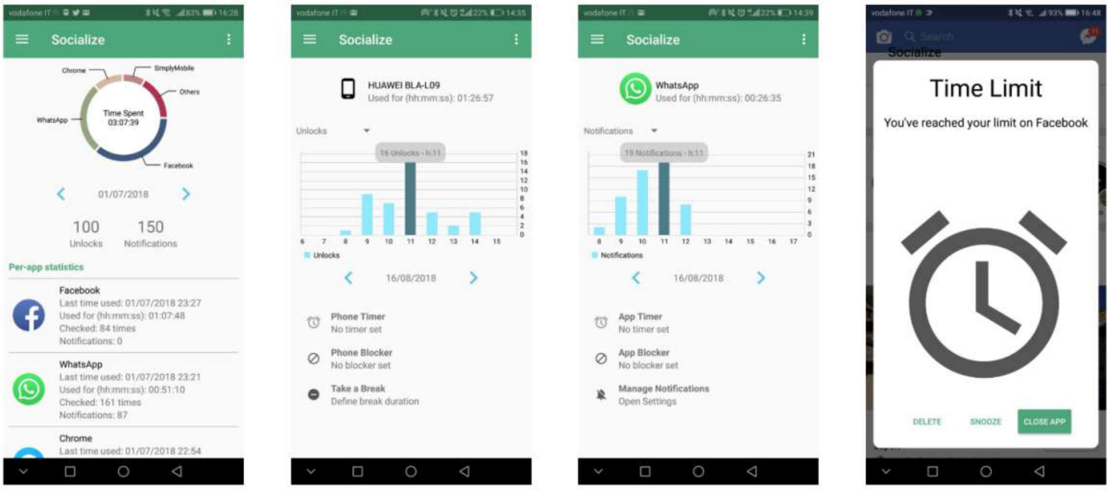
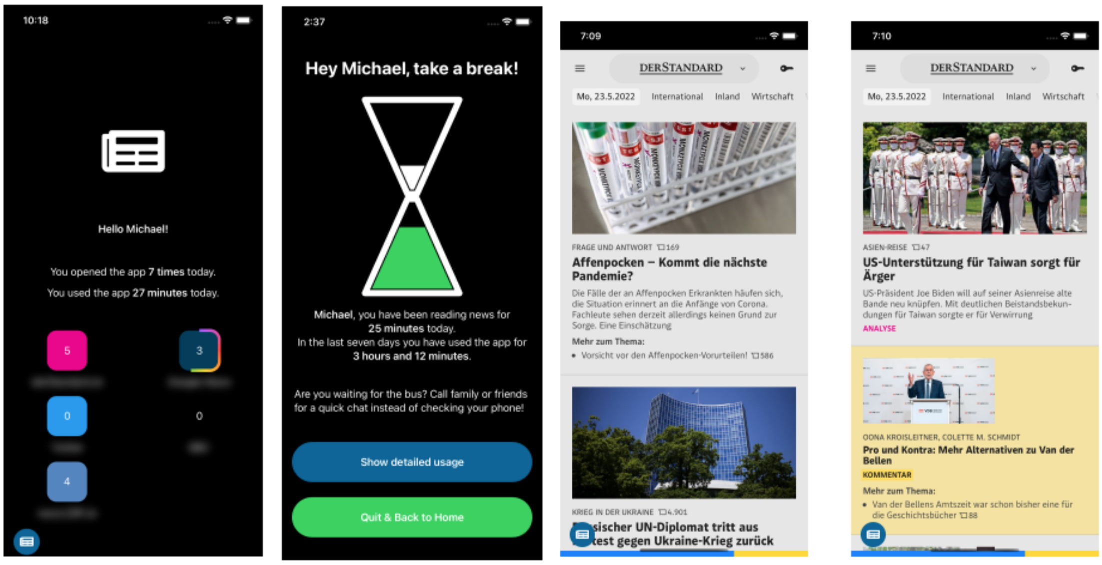


Figure 2.10: *Socialize* by Monge Rofarello & De Russis [MRDR19]

rarely used, and even the use of Socialize did not change how they perceive problematic phone use [MRDR19].

Purohit et al. [PH21] experimented with iOS **Shortcuts** application to create nudges to reduce the use of Instagram. Shortcuts allow users to program tasks without knowledge of any programming language and can thus theoretically be used with any app or content. The authors see the benefit of ensuring privacy and ethics by using this approach with built-in resources and mechanisms. The proposed intervention consisted of a Shortcut automation rule that would post a notification with the spent duration every minute after opening Instagram (for 10 minutes). A two-week evaluation study (one week of baseline, one week of intervention) revealed that the nudge significantly reduced the time participants ($n = 20$) spent on Instagram and was perceived as showing excellent usability.

Shahu et al. [SMWM22] proposed an application called **Nudgit** that tries to help with news consumption on smartphones with the help of digital nudges. The app was a host container for various news platforms and implemented three different mechanisms to reduce participants' news consumption. The effort to open a news portal (holding the associated portal button) was increased every time the news site had been opened within the same day (see Figure 2.11a) The app supported setting time limits for given news portals and displaying an overlay with statistics and recommendations for alternative activities upon reaching the goal (see Figure 2.11b). As a third nudge, Nudgit could filter news feeds by user-defined topics and block and remove unwanted articles (see Figure 2.11c) For evaluation, participants ($n = 5$) used the prototype over two weeks, with the first five days used as a baseline for news consumption — the nudges would be made available after a mid-study interview. A significant decrease in opening the app compared to the baseline versus the rest of the study was reported. Interviews revealed that participants further became more aware of their news consumption behaviour and that the nudges



(a) Nudgit's home screen (b) The reached limit's overlay. (c) Content filtering based on topics (COVID-19 and Ukrainian war)

Figure 2.11: Screenshots of *Nudgit* [SMWM22]

assisted the participants in reducing time spent with news platforms [SMWM22] There were no investigations regarding the long-term sustainability of the intervention, which is understandable, given the study scope (as a pilot) and size.

Again, Purohit et al. [PBS+23] employed iOS' **Shortcuts** app again to co-create nudges with participants for Instagram. The authors used the Shortcuts app to create an automation that would push a notification displaying how often Instagram had been opened when the social media app started (see Figure 2.12). Study participants ($n = 10$) would co-create an automation intervention for Instagram with a researcher after spending the first week capturing a baseline of usage. Instagram usage was tracked using another Shortcuts instruction logging timestamps to a local CSV file [PH21]. However, participants were told the research's purpose was to develop interventions for digital well-being that they and others would like to use. As they built the above-described nudge, the intervention's customisation was limited to the messaging displayed in the notification. Participants actively used the co-created intervention automation during the second week of the study. The authors reported a significant reduction in both Instagram usage time and the times Instagram was opened when comparing the baseline and intervention week. Additionally, participants felt to have a sense of agency and accomplishment. Perceived privacy, privacy concerns, and usefulness were reported as being very high, which would lead to sustainable behaviour. The authors attribute the success to a kind of IKEA effect [NMA12] as 'I designed it myself' fills users with more agency [PBS+23].

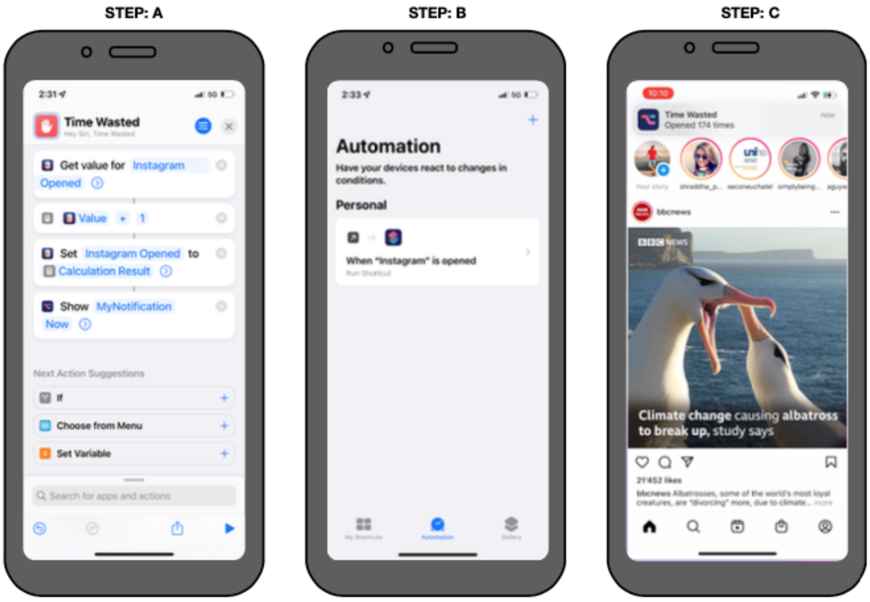


Figure 2.12: An intervention design by Purohit et al. using iOS' *Shortcuts* app [PBS+23]

2.5 Towards tangible devices and *Wearables*

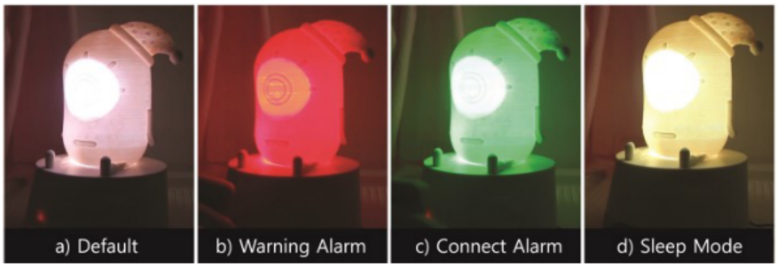
In recent years, wearable technology has found its way into consumer products and proposes to enhance our quality of life in ways a smartphone alone cannot achieve. These devices come in wristbands, watches, jewellery, glasses, clothing, and many other shapes. They can help their bearer to perform micro-tasks or sense and collect data around the clock [SHN+17].

Most prominent among Wearables are currently smartwatches, which have been termed the next killer product after the smartphone [JKC16, SHN+17] The history of body watches might have contributed to the first major lessons for wearable computing, as there are striking parallels in the evolution (minimisation, body positioning, affordability) [Mar02].

Wearables and smartwatches have been successfully used in various studies investigating behaviour-change interventions: promoting physical activity in sedentary lifestyle [SPA+15, SP13, MHM16]. Others have employed data from wearable sensors to identify symptoms of depression, bipolar disorder and schizophrenia [SDVJ+19].

Good examples of intervention methods are two smart lamps that try to reduce phone usage before bedtime and thus help to establish better sleep hygiene.

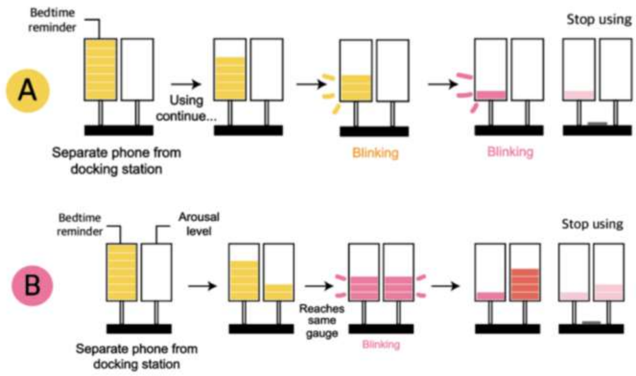
Lee et al. [LLKC17] proposed **D-Tox** a smart lamp that tries to help adolescent phone users reduce the frequency of smartphone usage during nighttime. The lamp would change its colour depending on various factors: if the phone is not placed in the bottom part of the lamp (see Figure 2.13b) after a defined amount of time, the lamp would



(a) *D-Tox* nudges [LLKC17]

(b) *D-Tox* [LLKC17]

(c) Prototype deployment [CLPK19]



(d) Prototype functionality description [CLPK19]

Figure 2.13: Smart lamps assisting building better sleep hygiene [LLKC17, CLPK19]

turn red to indicate it's time to stop using the phone. Otherwise, a 'cosy light colour' is enabled to help sleep better (see Figure 2.13a). The system also included gamification elements, such as a goal and reward system and a dedicated smartphone app for parents. Unfortunately, the prototype design was not evaluated by a study [LLKC17].

Choi et al. [CLPK19] proposed two prototype designs (see Figure 2.13d): Prototype *A* was a bedtime remember that would display the time until sleep. Prototype *B* reflected the phone usage (arousal bar) while showing the time till sleep. When a phone was put into the lamp's station, the arousal bar started to reduce again. A field study over three days revealed that participants ($n = 8$) preferred prototype A over B since B introduced more pressure to the participant, but in turn, it could reduce the time their phone was used before bed.

Both prototype lamps show how intervention design can change using different environments and involved devices.

Rudolph Schelling Webermann (RSW) proposed **TIME OFF!**[7] a watch with a digital time-out button for the phone as a design study. The watch would hold a simple digital display showing the time and a big red physical button (see Figure 2.14a). Pressing the

[7] RSW: TIME OFF!, `https://rudolphschellingwebermann.com/en/projects/off-` Accessed August 2023

(a) *TIME-OFF!* by *RSW* in normal watch mode.

(b) *TIME-OFF!* by *RSW* in time-out mode.

Figure 2.14: *TIME OFF!* a design study by *RSW*

stop-watch-like button would enable a time-out mode on the linked phone to help with distractions (see Figure 2.14b). Although only a design study with no fully functional prototype, the idea is simple and elegant.

Wearable technology opens up new design spaces for user interaction. It offers new opportunities for designers, developers and researchers that could be used to assist with the reduction of phone and application consumption.

It becomes apparent from the body of literature and available applications just very few applications currently use smartwatches to assist in digital detox activities or even represent functional prototypes

2.6 Regarding Health–monitoring and, activity and condition recognition with *Wearables*

Prevention and early detection of medical conditions are essential for promoting wellness. However, traditional clinical diagnostics often fail to detect early stages of issues because they are often only carried out in emergencies. Previous patient history is often partially unknown or sketchy, and practices are often complex to carry out in non-hospital environments or labs [MSKRJ17]. Hence, deploying devices in out-of-hospital environments could greatly assist individuals and medical practitioners, thanks to the comfort and daily care they provide to the wearer [IMD+21]. A broad range of clinical, research- and consumer-grade wearables already exist that could help revolutionise healthcare [DRS18, IMD+21].

Wearable systems can be employed to either try to prevent diseases by correcting unhealthy behaviours or provide responsive capabilities that detect medical conditions by monitoring and analysing biomedical signals over long periods [MSKRJ17]. Wearable systems used to monitor health are comprised of various miniature sensors, which are capable of measuring significant physiological parameters, e.g., heart rate, oxygen saturation, blood pressure, body and skin temperature, respiration rate, galvanic skin

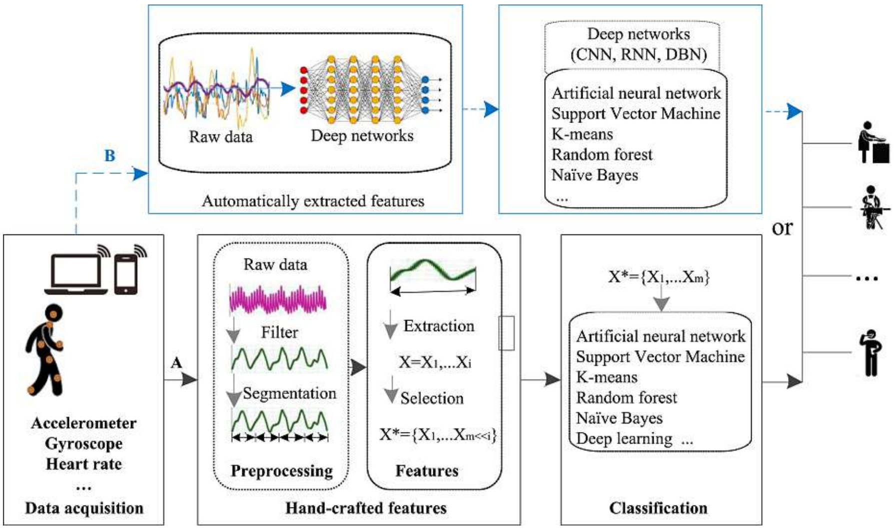


Figure 2.15: Learning procedures for activity recognition. [WCY19]

response [PB09, WCY19, CEF+12, MSKRJ17]. Other sensors might be added to capture environmental parameters, e.g. accelerometer, gyroscope, magnetometer, GPS, barometer, humidity, temperature, light sensor, etc. [MSKRJ17]

Raw sensor data can be processed and interpreted on the device and then communicated to other devices or cloud services for further processing or storage. Analysing raw sensor signals to recognise medical conditions or to detect physical activity like gait, steps, sleep cycles, etc., can occur on the device (e.g. [Udd19]) or as part of cloud services, depending on the complexity of the methods involved (see Figure 2.16 for an example system architecture). Recognition tasks can be solved with machine learning approaches [WCY19] (see Figure 2.15). Wang et al. distinguish the approaches between traditional recognition pipelines using handcrafted features and deep learning methods such as convolutional neural networks or recurrent neural networks [WCY19].

Wearables and smartphones offer promising opportunities to change physical activity behaviour. Systematic reviews have documented the positive effects of smartphones and wearables when promoting physical health [SAVL+16, GMvO+18, BWO+19]. However, often, studies only report short-term changes or only recruit healthy users.

Monge Roffarello and De Russis found in their systematic review that current digital self-control tools include a lack of theory and in general disregard ethical implications and issues. Secondly, the prevalent short-term evaluation leaves their performance in terms of building long-term sustainable habits up to speculation. Furthermore, the authors propose based on the review that future research might look into overcoming a limited perspective that focuses on technology overuse and self-monitoring tools and finally to find a way to deal with the business models of big-tech enterprises, that benefit from the continuous and frequent usage through their users and clients [MRDR23].

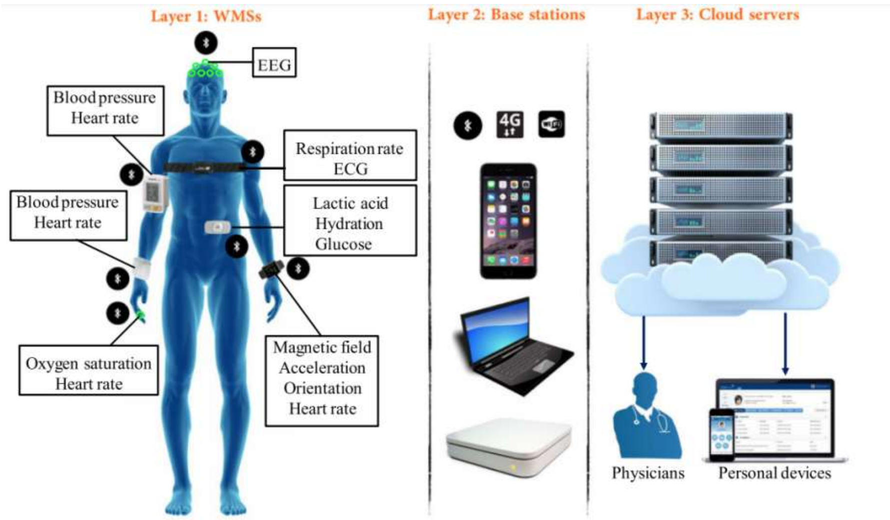


Figure 2.16: Architecture of a wearable medical sensors system [MSKRJ17]

2.6.1 Challenges

Wearable technology still faces numerous challenges, starting from regulatory issues and compliance to standards when being used in the health sector to power consumption and ethical concerns [FPE10, SHG22, SAACP18, PZ17, MC15]. Among these issues to users, the most relevant one probably is privacy since wearables capture data which is highly attributable to a user and provide intimate insights into a person's life and state of health or can be used for malicious activities [PZ17]

Another problem that often impacts user acceptance is the accuracy of sensor data or recognition results, as shown by Svenson et al. [SCT$^+$19] for detecting sleep cycles. The authors compared sleep measurements of the Fitbit Versa smartwatch reported with a validated single-channel EEG system for 20 participants and two weeks' worth of nightly data. The results showed no significant differences between time spent in bed and total sleep time, but the watch significantly over- and underestimated time spent in deep sleep and being awake. Despite this discrepancy, Svenson et al. found that this sleep tracker could still be helpful in measuring sleep duration in longitudinal studies [SCT$^+$19] with some limitations.

2.7 Combining wearable and phone usage data

Combining data from multiple devices or sources to recognise mental state, mental health, stress, and sleep quality is not new. Sano et al. [SP13, text] have collected accelerometer data and skin conductance data from a wearable sensor alongside mobile phone data to infer a measurement of general health, mood and stress.

The same Sano and different et al. [SPA$^+$15] collected data from wearable sensors, phone data along with self-reported measurements for sleep-score and personal traits tried to characterise factors which favour academic performance, sleep quality, stress and mental health.

Massar et al. have demonstrated that using wearable sleep tracking data, phone-based tappigraphy and self-reported provide redundant and complementary information about sleep behaviour [MCS+21]. Although tappigraphy does not involve phone-usage measurements such as screen time and which application participants used, it might be possible to estimate daily screen-time through deduction.

Arora et al. [ACB21, ACB22] have been using physical activity and sleep data from smartwatches and screen time attributes from smartwatches to assess sleep quality and behavioural health. The authors further propose methods to predict sleep quality based on these attributes with 91% accuracy.

2.8 Guidelines for designing digital detox applications

In addition to the overview of applications in Section 2.4 and 2.7, the following section presents recommendations and guidelines for designing and implementing digital interventions.

2.8.1 Nudge design

Roffarello & De Russis have summarised their recommendations for implementing digital detox applications with the following three suggestions [MRDR19].

Positive Reinforcements & attainable goals Positive reinforcements have been found to foster more sustainable long-term behaviour changes [APS97, CML09, MRDR19]. As Consolvo et al. put it, performing the desired behaviour should be rewarded. Still, when actual behaviour does not reflect the desired behaviour, it should not get either rewarded or punished, but the individual's interest should be sustained. There might be times when the individual is unable to achieve the desired behaviour due to various reasons, for which the intervention should not trigger guilt and risk abandonment [CML09].

Generating positive outcomes through choosing attainable goals within the boundaries of the individual's capabilities that lead to immediate or short-term satisfactory experiences has been additionally recommended in the area of building physical exercise habits [APS97]. It could also be applied to building digital behaviour habits [MRDR19].

Routines and trigger events Stawarz et al. [SCB14, SCB15] recommend habit-building applications to establish new desired habits (in their case, taking medication) as part of already established habits (e.g., after breakfast or brushing teeth). Timing and location make the routine more memorable, meaningful and reliable [SCB14].

Contextual awareness Finding the right timing to deliver the right nudges is essential [PH19]. The contextual awareness provided through smartphones and smartwatches could significantly help to determine digital, environmental and internal context [MRDR19, PH19]. However, choosing the right nudge is challenging if various choices are available. Recommending going for a walk when the individual had already finished a physical activity a few hours earlier might lead to feelings of frustration and loss of credibility and consequently increase the risk of abandonment.

Wearable devices could provide additional information to identify context and time interventions more precisely.

Purohit et al. proposed the following guidelines for nudge design [PBS+23]:

Transparency Gold et al. verified that transparent behavioural interventions were considered more acceptable than opaque ones. A transparent intervention is when the mechanism that is used to influence behaviour is identifiable, or the mechanism is being disclosed to participants [GLAO23]

Support Opting-out Receivers of nudges should have the choice to follow or unfollow a nudge. If reflective or deliberative processes of decision-makers are ignored, their autonomy is lost [PBS+23]. Ignoring the individual's independence comes close to executing paternalism, which also goes against Thaler&Sunstein's recommendation [TS08].

Nudge for good As proposed by Thaler in his criticism of 'sludge', using nudges should ultimately have the best interests of the nudge-receiver in mind [Tha18]. This extends Thaler&Sunstein's definition of nudges, saying that nudge-receivers should not be put under financial stress by following the nudge [TS08] to new dimensions. Münscher et al. argue that putting higher than marginal efforts effectively changes the choice architecture since extensive costs do not justify behaviour change on rational grounds [MVS16].

2.8.2 Application design

Consolvo et al. have built on the design of Jafarinaimi et al. [JFHZ05] to propose a design of eight strategies for lifestyle behaviour change technologies [CML09]:

Abstract & reflective Applications should abstain from presenting raw or explicit data to help with impression management [Gof16] when presenting in various situations (social), thus increasing the opportunity for reflection.

The degree of abstraction required depends on the individual's preferences, at least in the author's opinion. However, reviewing raw data to aid transparency should still be an option.

Non-obtrusive To collect and present data, making it available without unnecessary interruptions or calling attention when it's not required.

In this regard, smartphones and watches provide the best examples of potentially unobtrusive technology, as they are commonly available and accompany individuals daily to the point where people feel naked without their phones. However, the timing and design of notifications should be considered with great effort, as already discussed with the timing of nudges.

Public Personal data should be represented to make it possible to share it in a public space without feeling uncomfortable. There is always the chance of unwanted glances, especially when technology is used daily and in public.

Aesthetic Display design and accompanying devices are meant to be used throughout everyday activities over a long time. Consequently, these items must sustain curiosity, be comfortable if worn, and match the user's style.

Positive Again (see Section 2.8.1), positive reinforcements are recommended to support long-lasting behaviour changes.

Controllable Consolvo et al. recommend allowing manipulation of data to allow data to reflect behaviour deemed suitable by the user [CML09]. This can become challenging when behaviour is automatically inferred, as accuracy plays a vital role in sustaining the application's credibility to the user. In the author's opinion, this is closely linked to transparency, chances of opting out and context awareness discussed above (see Section 2.8.1)

Trending/Historical Providing access to historical behaviour might make patterns of singular bad choices more readily apparent. Applications encouraging behaviour change should offer the possibility to reflect on past behaviour in relation to reaching a self-defined goal.

Comprehensive Accounting for a range of behaviours and using all available means technology provides to sense or monitor. Consolvo et al. provide the example of recommending physical activity with their implementation, which only used a pedometer to track walking activities. Consequently, users get discouraged when riding a bike and do not receive credit for performing physical activity.

2.9 Going forward

Looking at the history and the current state of digital detox applications to support the reduction of screen time (see Section 2.4) a few things have become apparent:

Firstly, a deep integration into the operating systems is required to enforce certain capabilities to limit from a technical perspective. Given the recent development in strengthened security in mobile devices' operating systems, getting access to the means to control processes often means breaching security measures. Furthermore, strict or inflexible interventions have led to user frustration in the past and motivated participants to either find loopholes or abandon the experiment. Most of the tools available, are confined to the smartphone itself and rarely make use of technology beyond this plane.

Secondly, from a sustainable habit perspective, participants tended to return to old usage patterns after the intervention experiments. Yet this behaviour has been reported through catch-up interviews by the participants. Interesting nudging mechanics are available but their longevity in upholding users' motivation dwindles over time. Consequently, finding the balance between hard and soft interventions poses a challenge in overcoming technical implementation limits and finding the sweet spot between usability and hindrance to support sustainable habits. This highlights the need for longitudinal studies and long-living intervention mechanics to keep participants interested and motivated for a longer duration. The literature does not provide consent on what is a feasible approach to deliver this to participants on a larger scale. Given the outcomes of the presented studies, one can assume it highly depends on the individual's attitude and motivation towards reducing phone, applications or social media usage when it comes to having a long-term meaningful impact.

The currently available state in the literature has highlighted the chance for approaches moving beyond the smartphone and using additional context and interaction potentials provided by such means.

CHAPTER 3

Methods

Having identified the gap in digital detox applications that employ smartwatches in both the available literature and Google's and Apple's app stores, the logical next step was to design and implement a prototype. A proof of concept intervention design and application shall help answer research question **RQ 1**: "How might smartwatches aid with digital detox activities"?

A preliminary user study was conducted to gather insights into how users perceive their phone usage and how they manage their screen time (see Section 3.1). With reference implementations (as presented in Section 2.4) and design recommendations (see Section 2.8), a possible intervention design is presented in Section 3.2. Sections 3.3 and 3.4 describe technical design considerations for a possible prototype, whereas Sections 3.5 and 3.6 describe Android and WearOS implementations of the proposed prototype implementation.

To evaluate the prototype, a user study was conducted (as described in Section 3.9), during which quantitative data was captured to measure the effect of the prototype on phone usage. An interview was additionally conducted with participants to capture their personal experiences. To conclude this chapter, Section 3.9.3 describes data processing for further analysis. The evaluation of results (Section 4), in turn, will then help answer the remaining research questions (**RQ 1.1**, **RQ 1.2**, **RQ 1.3**, **RQ 2** and **RQ 3**).

3.1 Preliminary interviews

Six Participants ($n = 6, f = 4, m = 2$, age average=29.8, $\sigma = 6.96$) were asked about their phone habits, usage awareness, their strategies to limit time spent on the phone, their experiences with wearables and their willingness to wear a prototype with our proposed features.

(a) Standard watch face of the *Google* Pixel 2 watch[1].

(b) Draft of the prototype "watch face".

Figure 3.1: Watch face example and early sketch

Most users claim they are aware of their phone usage. Nonetheless, some participants had sudden realisations during the interview while discussing their daily routines and were surprised when they saw their usage statistics. However, others are aware of the monitoring apps but have chosen to either ignore reports or not actively seek out how to change their behaviour.

Participants have reported having reverted to uninstalling social media applications for a short duration of time to either focus or try to detox. Others have been putting more effort into being able to use certain apps by using browser versions instead of native applications. But eventually re-installed or reverted to the applications. Regarding wearables, most participants have had experiences using fitness trackers and would test a prototype application. However, concerns were reported regarding wearing watches due to feeling restricted by the watch and adding yet another piece of technology to a digitally overloaded life.

Two participants have described that they would feel helpless without their phones in terms of handling some of their everyday activities, and this had already interfered with past digital detox experiences.

Due to the small number of participants, the findings were not conclusive. However, performing these interviews was useful in building rapport and seeing different perspectives on phone usage.

3.2 Intervention Design

Based on applications found during the literature review and by reviewing available reference applications, the following two intervention designs were designed, following roughly the design of Okeke et al. [OSDE18].

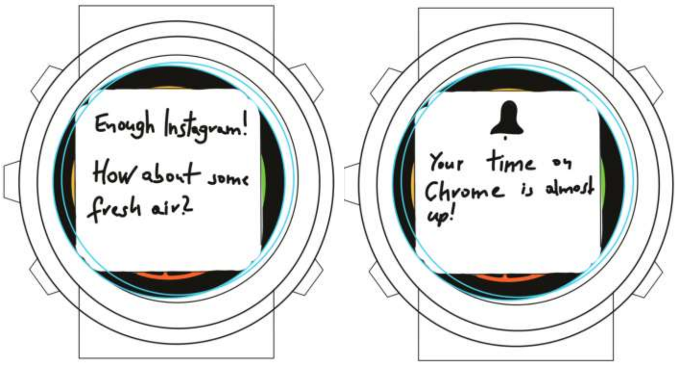


Figure 3.2: Early draft of the “alarm nudge”-intervention.

3.2.1 Watch face

Displaying data on the watch’s face is a simple and effective way to provide information to the watch’s owner without causing much obstruction and distraction (see Figure 3.1a as an example). Similar applications can be found with almost every watch that is capable of processing data, whether it’s in the medical (e.g. blood-sugar measurements, heart-rate monitoring), sports or daily-live (notification or navigation) or entertainment (games) field.

Watch faces have specific characteristics that make them problematic for showing sensible information: watches are worn openly and are supposed to be readable with a glance. However, in the case of non-standard watch interfaces, such as a heavily personalised watch face, only a knowledgeable viewer can interpret the data.

Thus, following the design goals proposed by Jafarinaimi et al. [JFHZ05] when developing Breakway (a smart sculpture displayed in a public space motivating breaks when sitting too long) was a sensible way to approach the design for the watch face (also see Section 2.8):

The watch face serves as near real-time feedback for the user at all times of the day, even outside of phone interaction, thus hopefully raising usage awareness. Figure 3.1b shows an early draft of this intervention.

3.2.2 Alarm nudge

The design was lent from Okeke et al. [OSDE18]’s implementation. Consequently, vibration is employed as a feedback nudge on the watch, triggered when a defined limit of daily screen time is reached. This intervention should require users to lift their attention from the phone and reflect upon their usage.

The vibration can be dismissed only on the watch; thus, focus on the phone is shortly drawn away from the phone. If participants continue using a limited application after

[1] Google Pixel 2, `https://store.google.com/gb/product/pixel_watch_2`–Accessed January 2024

the daily goal, the nudge will be triggered again.

The vibration is accompanied by a message displayed on the watch, saying:

> You have been using "<*application name*>" longer than you intended.

The wording was intended to be factual and neutral.

Other than Okeke et al's limits for Facebook, participants can create and define the limit for whatever application they feel they should get some distance from [OSDE18]. Figure 3.2 shows very early drafts of the visual feedback of the alarm nudge intervention.

3.3 User interface design considerations

Before and during application development, certain design-related aspects had to be evaluated and decided on, which shall be presented in the following section.

Watch face first

The watch face represents the element with the most constraints imposed upon it, as recommended by Google's WearOS design principles[2].

- There is limited space available
- The face needs to show relevant information together with showing the time of day without being cluttered
- Simplicity and readability are key

Hence, since the face can be considered the central point of interaction, it makes sense to transfer parts of the design to the phone application. This approach aligns very well with Consolvo's design recommendation [CML09] as it simultaneously covers ticks for the proposed abstract, non-obstructive, public and aesthetic themes (see Section 2.8.2). An example of how this is achieved in *Watchful* can be seen in Figures 3.4a and 3.8a.

With the circular shape dominating the initial touch-points of the watch and phone app's user interfaces, re-using the round element elsewhere adds further to design consistency criteria. Hence we see the shape reappear as the time-doughnut (see Figure 3.5c) or the alarm-input-slider (see Figure 3.4c).

[2] WearOS design principles, `https://developer.android.com/design/ui/wear/guides/foundations/design-principles`–Accessed August 2023

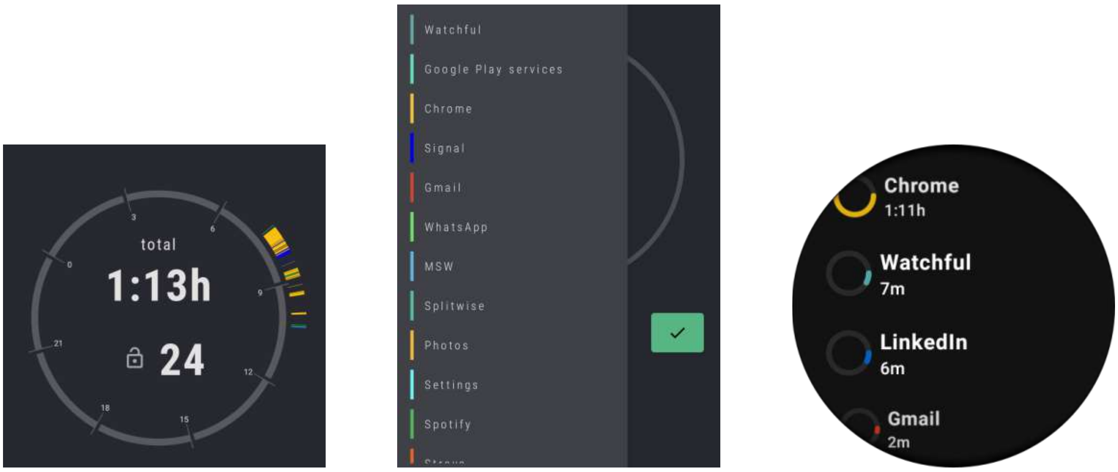


Figure 3.3: Colour consistency per application across views and devices.

Colour consistency

Colours carry information. In this case, it stands for which application was used in overviews and visualisations and when or for how long an application was used. Consequently, colour mappings should never change and must be consistent across all user interfaces on the phone and watch. This behaviour is achieved by keeping a registry of all installed (current and passed) applications with a colour code representing the primary colour of the application's icon. Figure 3.3 shows examples of different views available across different devices using the same application and colour mapping schema.

Colour is an additional attribute inside all synchronisation or alarm messages to uphold consistency across all involved devices.

3.4 Platform selection for phone and watch

Mobile phone operating systems used to be a vibrant and thriving area until just Google's Android and Apple's iOS remained as relevant players on the market (market shares: Android: ~ 68%, iOS: ~ 30%)[3].

APIs to access information about application usage were available on both operating systems at the time of planning development for this thesis:

- iOS:*Screen Time API*[4] available since *iOS15* (released 2021).

[3] Mobile Operating System Market Share Worldwide, `https://gs.statcounter.com/os-market-share/mobile/worldwide`–Accessed May 2023

[4] Screen Time API, `https://developer.apple.com/documentation/DeviceActivity`–Accessed May 2023

- Android: *UsageStatsManager*[5] available since *Android 5. — ollipop* (released 2014).

However, since development experience was already available for Android, the decision was easy to make in favour of Google's mobile phone operating system and ecosystem.

Having settled on the phone platform, left the selection of the smartwatch platform with a few promising options to choose from:

- **watchOS:** Apple's watch platform does not support full-feature integration with Android phones and hence was not further considered a development platform.
- **Garmin:** Mainly known for sports watches and navigation devices, Garmin provides API and SDKs.
- **Suunto** provides the SuuntoPlus partner program to which potential developers need to apply first.
- **Tizen:** Developed by Samsung and has been integrated gradually into Wear Os since 2021.
- **Wear OS:** Google's Android for wearable devices. Supports easy integration with native Android applications through a shared code-base and ecosystem.

The idea of using a sports watch without too many smart features as a platform was the obvious choice to be used as a companion for digital detoxification activities. However, a proof-of-concept design based on Garmin's platform failed to deliver the vision of this work due to limitations with background communication between the phone and watch to synchronise the latest usage data.

Consequently, switching to Wear OS was considered the logical next step. Therefore the only available sports model available at the time (spring/summer 2022) was chosen, which was the Suunto 7 running Wear OS 2.5. Because Google Fit came preinstalled on this watch model, the selection for the back-end for health-related data and activity tracking was fixed. With all software components originating from the same source (Google), integration and inter-play are expected to be well documented and tested.

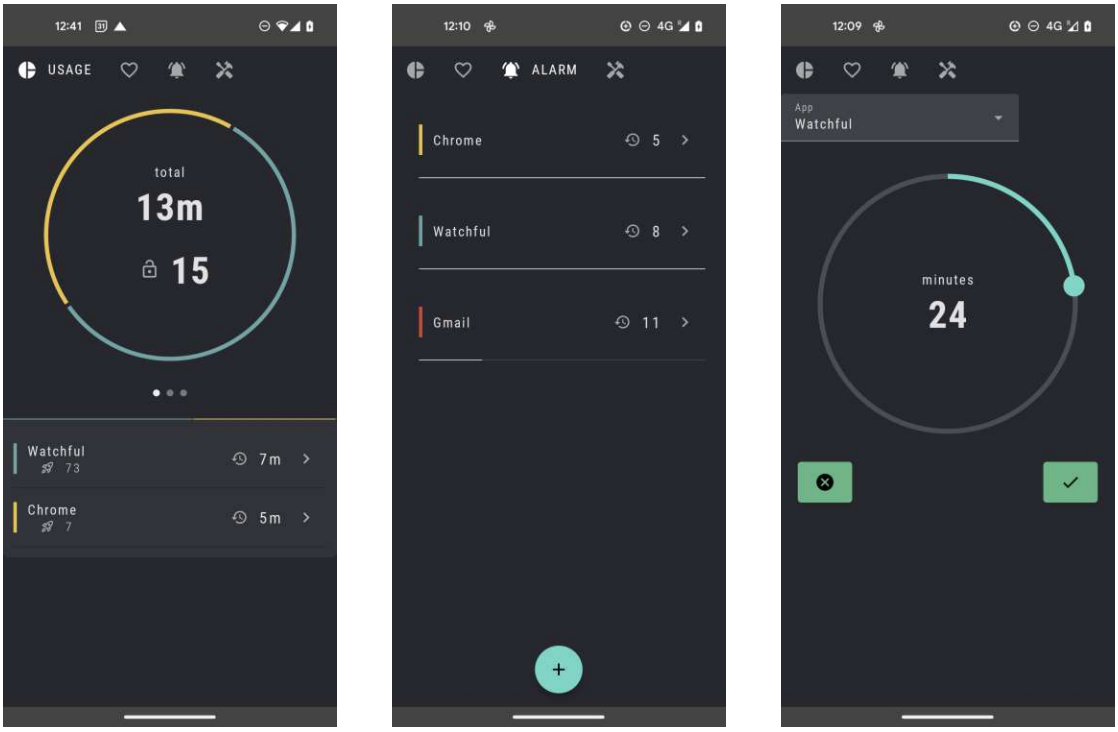
3.5 *Watchful* phone application

3.5.1 Technical design

To provide the best possible user and developer experience with the available resources, Google's recommendations for modern Android development[6] got considered and followed throughout all components.

[5] UsageStatsManager, `https://developer.android.com/reference/android/app/usage/UsageStatsManager`–Accessed May 2023

[6] Modern Android Development, `https://developer.android.com/modern-android-development`–Accessed July 2023



(a) Overview view, identical state to Figure 3.8a

(b) Overview of active alarms and new alarm button.

(c) Configuring alarms with a circular slider element.

Figure 3.4: User interface of *Watchful*

The application is written in *Kotlin* and follows a single-activity architecture with modular components for respective features. The user interface is based on Rally, a material design study by Google[7] and is implemented in Jetpack Compose.

For data persistence on the participant's phone, *Room* was used as an abstraction layer over SQLite, the database standard for Android. App and phone usage events, which got periodically scraped from the Android system's own StatsManager API, were inserted into that database and played the backend role for the user interface. However, since usage data is sourced from both the database and the StatsManager, a unified representation was introduced based on *Kotlin dataframe*[8], which also allowed simple statistical calculations and to implement graphical representations against the data frame as interface. Custom visualisations to plot bar and line graphs were added, and non-standard input elements like the slider to configure alarm limits (see Figure 3.4c) were also implemented in Compose.

[7] Rally Material Design, `https://m2.material.io/design/material-studies/rally.htm` and `https://github.com/android/compose-samples/tree/main/Rally`–Accessed July 2023

[8] Kotlin Dataframe, `https://github.com/Kotlin/dataframe`–Accessed July 2023

3.5.2 Data visualisation design and user interface implementation

Several visualisations are available as part of the phone application, which can be separated by the topic of data they primarily display: usage data or health data.

Usage data

Usage data originates from Android's *UsageStatsManager* APIs and is used to populate various screens.

Daily overview The overview screen is dominated by a doughnut chart that shows the screen time of applications, and the count of phone unlocks since midnight (Figure 3.5a). Its design is very similar to the dashboard provided by *Digital Wellbeing* (see Figure 2.7a). Below the chart is a list of all applications a user has been using today. This overview can be used to navigate to the respective app-specific view.

Daily overview over time The doughnut chart offers a second mode which a swipe gesture can activate (see Figure 3.5b) It shows a clock-like representation displaying the past 24 hours with coloured segments that highlight the time of day during which the application was used. This view makes it easy to spot times of more extended usage.

App specific overview The app-specific view again shows the past 24 hours' usage of the selected application's usage in the clock-like representation (see Figure 3.5c). In case alarms had been triggered on that day for the selected app, the events would get displayed in the clock display (marks in white on the inside of the clock not to cover up usage segments) and a list of timestamps at the bottom of the page.

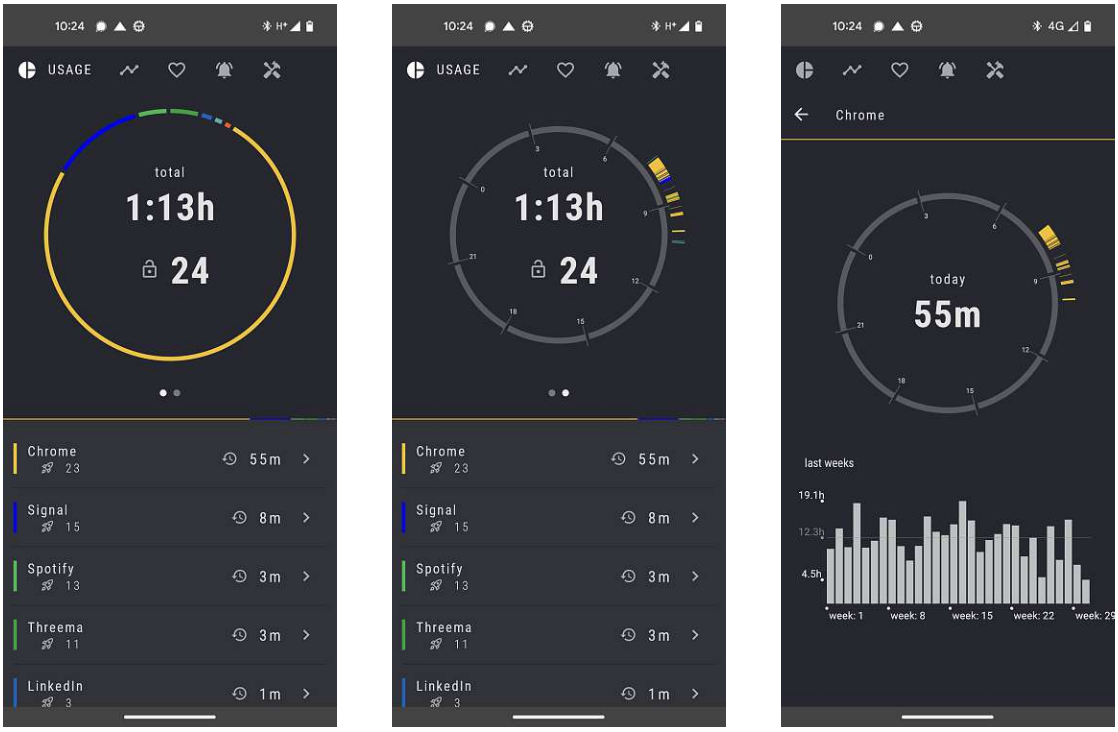
Below the clock is a bar chart with the weekly total screen-time usage over the last weeks (as far back as data has been scraped by *Watchful*).

Overview of historic data A separate view (the second icon from the left in the navigation bar) would display the weekly screen time totals as a bar chart, similar to the historical view for specific applications.

Health data

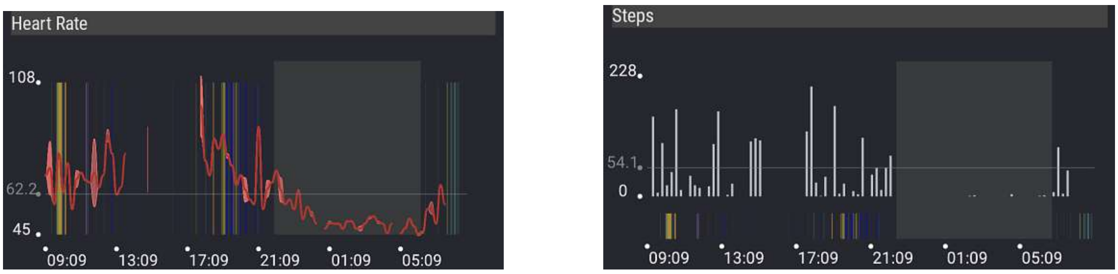
Health data displayed in the phone application would be queried from Google Fit, which generally displays the last 24 hours of data.

Heart rate Heart rate is aggregated in 5-minute time-buckets by Google Fit and displayed as a line graph (see Figure 3.6a; average of that 5-minute time window in dark red; min and max spreads in bright red) App usage (colourful bars) and sleep data (large grey block) are also incorporated into the visualisation.



(a) Total daily usage overview (b) Daily historic overview (c) App-specific view

Figure 3.5: User interface of *Watchful*



(a) heart-rate data with phone usage and sleep data (in grey).

(b) step data with phone usage and sleep data (in grey).

Figure 3.6: Health data visualisations for the last 24 hours.

Steps data Similarly to the heart-rate data, step data is displayed as a bar chart in 5-minute aggregation, with usage data displayed on a separate axis (see Figure 3.6b).

Sleep data Sleep activity data is incorporated in the heart rate and step visualisations as large grey blocks that show bed and wake time (as calculated by Suunto's employed metrics). Although sleep cycles are available, they are not displayed in the visualisations for simplicity reasons and for apparent lack of accuracy (see Section 2.6).

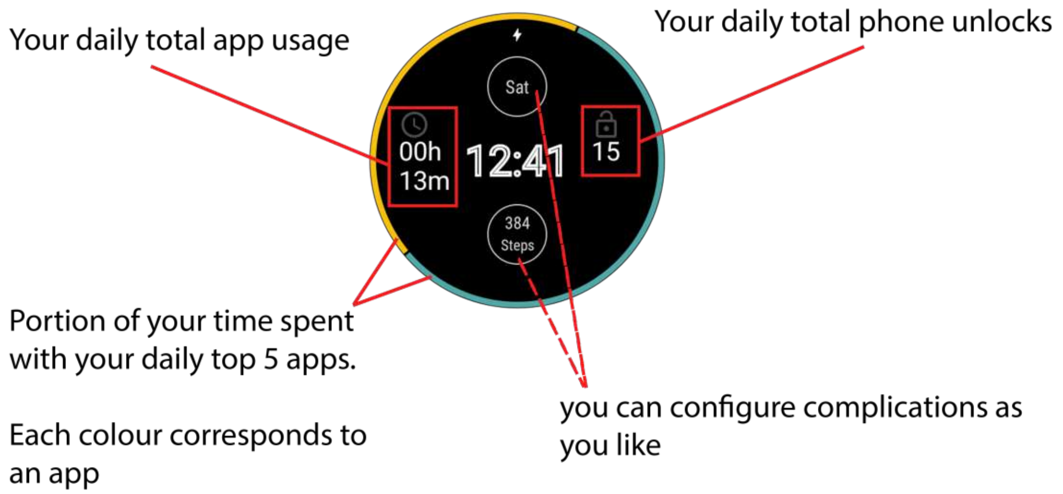


Figure 3.7: Explanation of the available fields of the *Watchful* watch face that was supplied to the participants.

3.6 *Watchful* watch application

3.6.1 Technical design

Like the phone application, the watch companion app was written in *Kotlin* and Jetpack Compose. From an architectural point of view, the app consists of two components that all contribute to the experience: A backend *app* providing the communication and persistence layer, which is based on protocol buffers[9]. Furthermore, separate activities were developed to handle the alarm (see Figure 3.8c) and provide a more detailed overview. Finally, the *watch-face* (see Figure 3.8a) implementation extends a *WatchFace Service*, which is responsible for displaying the current state stored in the persistence layer. See Figure 3.7 for an explanation of the provided information on the watch-face.

Nudges on the watch are triggered through the phone application when a limit has been reached for the first time. In case the participant continues using the app, an alarm gets triggered after a 10-second threshold period.

[9] protobuf,https://github.com/protocolbuffers/protobuf–Accessed July 2023,

(a) *Watchful*'s watch face — identical state as Figure 3.4a

(b) *Watchful*'s list activity on the watch.

(c) *Watchful*'s test alarm activity on the watch

Figure 3.8: *Watchful*'s watch application screens

3.6.2 User interface implementation

Watch face

The watch face would display the top five used applications (by screen time) and a 6^{th} 'other' category as a doughnut chart (just as being used in the phone application as the central element, see Figure 3.4a). On the left side of the time-of-day, the total daily screen time was displayed (in hours and minutes), and on the right side, the total number of phone unlocks was displayed. Two complication fields above and below time-of-day could provide customisation options (see Figure 3.7). They could be used to display various other data, e.g., date, battery charge, weather data, etc.

In light of the design guidelines by Jafarinaimi et al. [JFHZ05] discussed in Section 3.2.1, the implementation fulfils them partially.

1. **Abstract**: We show only aggregated application data and usage totals.
2. **Non-intrusive**: Watches and their faces are unobtrusive by design.
3. **Public**: Data is only glance-able only by the informed user as colour mappings are unique for each user. However, total usage time is problematic but is required.
4. **Aesthetic**: Customisation options are available through complications and colouring schemes.

The watch application features an activity besides the watch face, which can be used to see a more detailed data version (see Figure 3.8b) in the form of a list featuring at most six entries (top 5 applications and an *other* bucket). Each list entry features the application name, the daily screen time and a doughnut chart of the total daily screen time percentage. The alarm view (see Figure 3.8c) would be dominated by the colour of the application that was responsible for triggering the alarm (logo and OK-button colour).

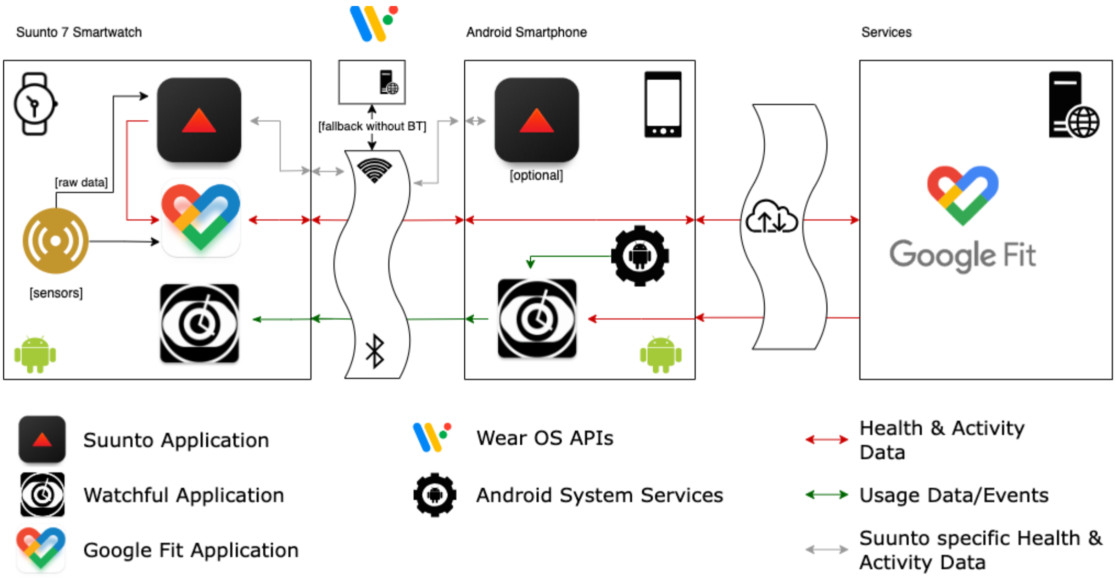


Figure 3.9: Data flow between all components.

3.7 *Watchful*'s components communication

Figure 3.9 shows the data flow between the most relevant components of this prototype design. The WearOS data layer displayed in the diagram needs some further explanation as, by default, communication between the phone and wearable devices is handled via Bluetooth. In case Bluetooth is disabled or not available, communication falls back over a Google Service via WiFi (or LTE)[10].

Application or health data never left the phone or the watch unless participants exported their usage data to provide researchers with their data sample for further analysis after completing the study.

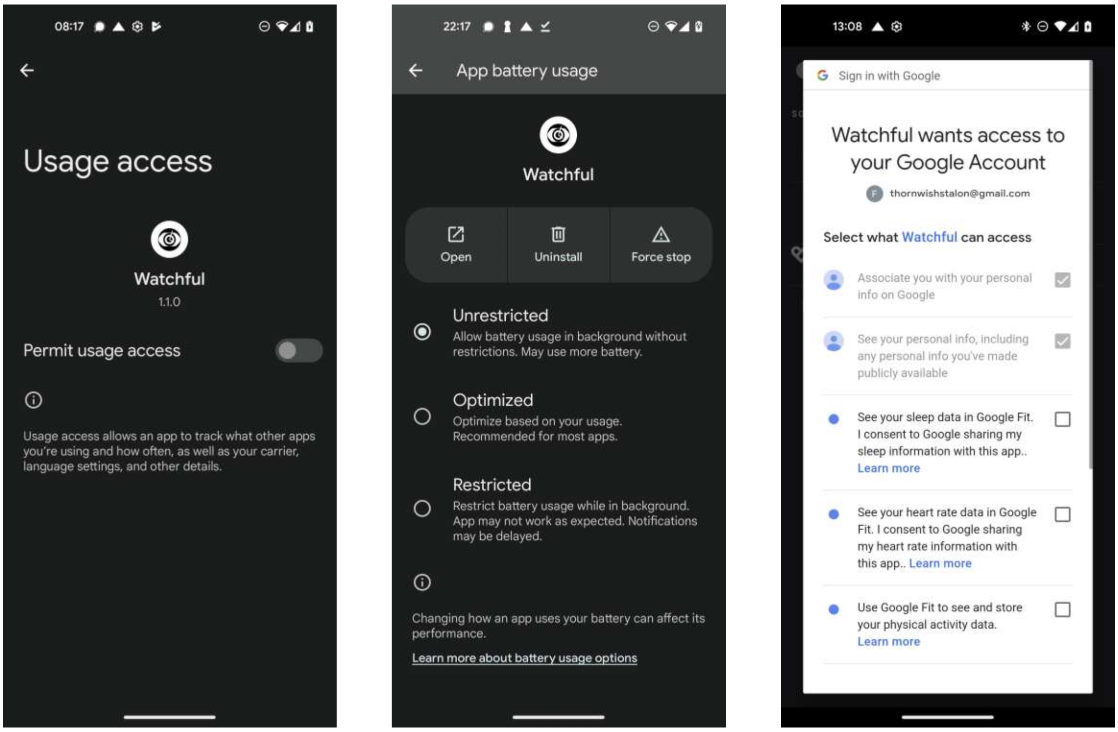
Communication with the companion watch application (see Section 3.6) was taken care of by the WearOS APIs (e.g. data layer) and used to send the latest daily usage statistics or trigger nudges (see Figure 3.9).

In order to reduce broadband and traffic and to preserve the battery of the watch and phone, only necessary update messages were transferred, which means that inactivity on the phone would result in zero transmissions on the usage state.

3.8 Application distribution and installation instructions

Phone and watch applications were uploaded in Google's *Play Store* in a *testing* state to make distribution to participants easier. Participants got invited to the tester group

[10]WearOS data layer, `https://developer.android.com/training/wearables/data/data-layer`–Accessed July 2023



(a) Permission for usage statistics

(b) Unrestricted background processes

(c) Permission to read Google Fit data

Figure 3.10: Granting required permissions for *Watchful*

and were thus able to download and install the dedicated apps on their phones and the provided smartwatch.

Additional setup steps were part of the onboarding meeting, which involved setting up the Google Fit and Suunto applications on the watch to be able to get access to heart rate, steps and sleep data, as well as providing Watchful with the participant's permission to read their health data (see Figure 3.10c) and disable battery saving policies for the app as this would interfere with scheduling background tasks (see Figure 3.10b). The most crucial step was to grant the app permission to retrieve phone usage statistics (see Figure 3.10a). The last step needs to be done manually, as starting with Android 13, permission to access usage statistics has been promoted to special permission status that can't be requested programmatically.

Despite being cumbersome during the setup, these steps are required to provide basic functionality and, in addition, should add to data transparency and governance as they leave the user in control.

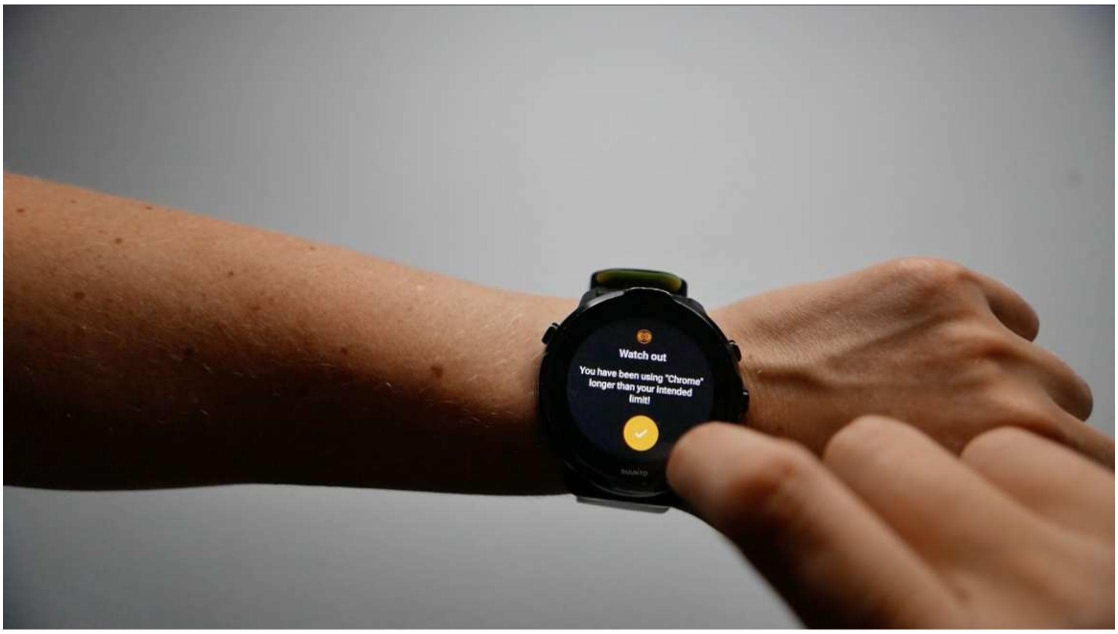


Figure 3.11: *Watchful*'s watch alarm in action

3.9 User study

3.9.1 Study Design

Despite the identified need for longitudinal studies in this regard (see Section 2.9) this study represents a pilot study to evaluate the proposed prototype design, hence the relatively short duration of two weeks.

Computer-science students ($n = 9(f = 2, m = 7)$) were recruited from a pool of bachelor course's participants. The selection criteria were based on their willingness to test digital detox for themselves and owning an Android smartphone. As compensation, participating students received bonus points that counted towards their final grade for a course. Students willing to participate had to read an information sheet and sign a consent form before partaking in the study.
See Figure 3.12 for a graphic overview of the user study design.

Smartwatches were provided to the participants during the study, and they were guided through the setup process as part of the briefing meeting (always scheduled on a Monday). Instructions for the first week of the study consisted of wearing the watch as much as possible (also while sleeping) to get used to it and the watch face, just letting the application observe the current behaviour. At the end of the first week, an email was sent to the participants as a reminder to start using the nudging functionality (see Figure 3.11 for an example) for selected applications during the second part of the study. Along with the mail came a short survey to catch first impressions and address technical difficulties and general questions.

Study phases

1. **Benchmark phase**:
   - 7 days prior to the beginning of the study.
   - extraction of screen-time only as a reference.
2. **Passive phase**:
   - first 7 days.
   - participants should accommodate themselves to the watch.
   - only monitoring of screen-time through *Watchful* applications
   - **no alarm intervention**
3. **Active phase**:
   - last 7 days
   - continuation of monitoring
   - **active alarm intervention**

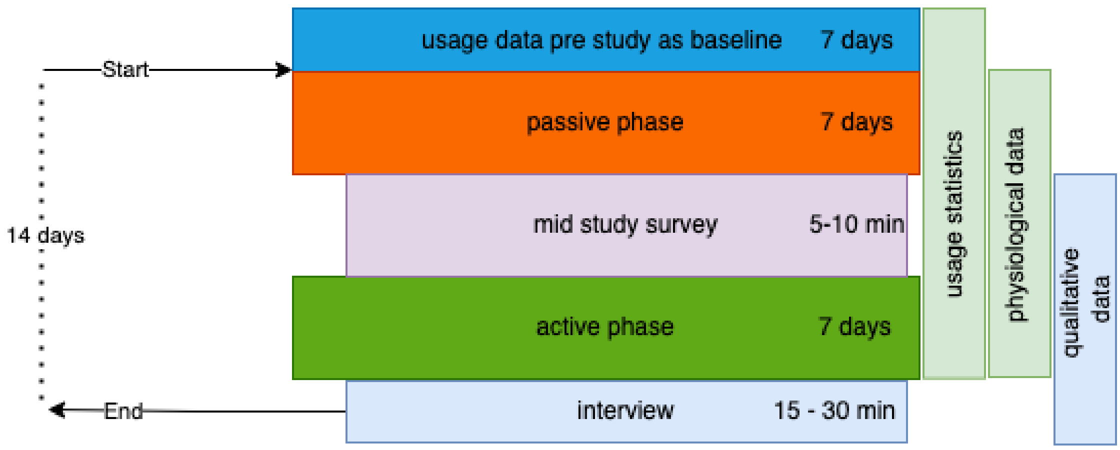


Figure 3.12: User study design

After completing the two weeks of trials, participants were asked to return to the lab to hand in an export of their data and hardware and be interviewed about their experiences. Audio recordings of the interview were captured using a handheld recorder.

Due to the limited number of watch devices available, just three participants could participate in parallel. This approach proved helpful, especially during the first iteration, in further streamlining the briefing and setup processes and addressing technical problems. For example, manufacturer-specific characteristics of some Android distributions needed to be accounted for so as not to let system processes or unique UI widgets skew the usage statistics and derange user experiences. Software patches could be prepared and distributed to participants promptly through Android's *Play Store* to address minor issues.

In contrast to other studies [OSDE18, KJKL19, PBS+23, PBH20], the usage in the week before starting the study was considered as the baseline reference to compare the next two weeks.

3.9.2 Measurement

With access to the participant's phone usage history, it was possible to extract the following metrics to determine the participant's behaviour for the duration of the study and at least seven days before the start of the study:

- start and end timestamp of applications running in the foreground
- timestamps of phone unlock events
- timestamps and application origin of the alarm
- timestamps, application and nature of alarm-settings changes (creation, deletion or update of limits)

Thanks to the integration of *Google*'s Google Fit[11] service into the watch, additional physical metrics were available to be used in the phone-application visualisations and post-study analysis:

- heart rate (time series data)
- steps (time series data)
- sleep activities (begin, end & sleep cycle)

3.9.3 Data processing

Data export

Participants were asked to provide us with a dump of their data at the end of the study through an export feature built into the phone application. The application gathers data from a local database, and remote Google Fit services write domain-specific CSV files marked with a UUID on the phone's internal storage. As a last step, a zip-archive was created to make sharing with the authors through a channel of the participant's choice more convenient.

[11] Google Fit, `https://www.google.com/fit/`–Accessed May 2023

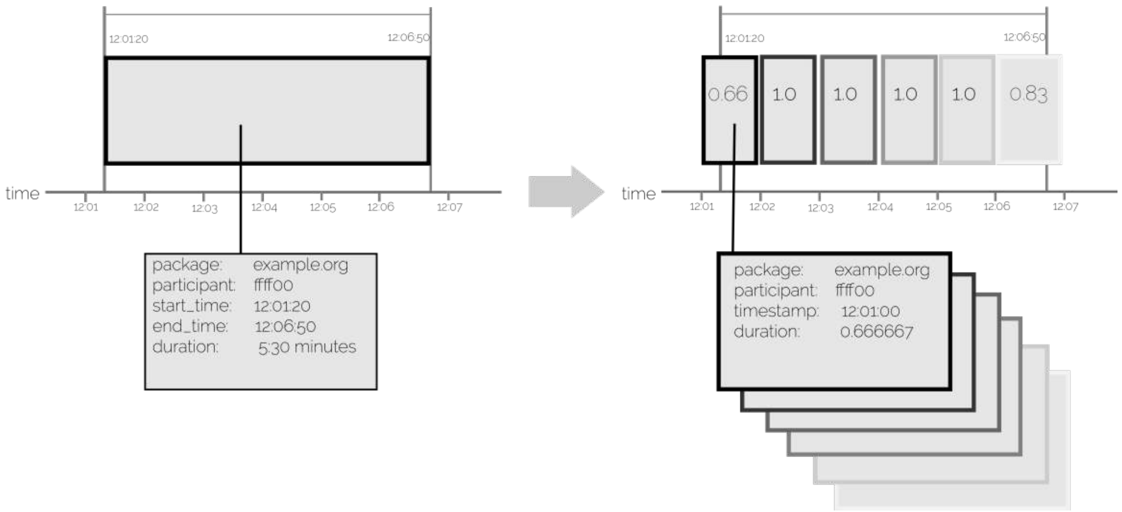


Figure 3.13: Transforming descriptive event data to data points for easier analysis.

Data transformation

An import routine fed the zip-archived files into a dedicated timescaleDB[12], database — running locally on the author's notebook — as a unified source for all the study's data.

Due to timescaleDB providing powerful tools to interact with time series data, event data describing ongoing user interaction with the application in the foreground got transformed from a descriptive representation into time-series data points in **one-minute resolution**. Figure 3.13 shows the data transformation based on an example event, resulting in six data points with a duration value that is fractional of the event duration at the given timestamp of the full minute.

Consequently, preparing data for statistical analysis for arbitrary time-frames and buckets with optional gap-filling functionality through this transformation was simplified.

Data exploration and visualisation

For an initial exploration of data that was then available in a unified form in a database, the visualisation tool *Grafana*[13] was used through dashboard with custom view implementations (see Figures 3.14a and 3.14b).

Interviews recordings

The recorded interviews were processed using an online transcription service called *sonix.ai*[14] to create a preliminary transcript that got manually checked for accuracy by

[12]TimescaleDB,`https://www.timescale.com/`–Accessed May 2023
[13]Grafana Labs, `https://grafana.com/grafana/`–Accessed November 2023
[14]sonix.ai, `https://sonix.ai/`–Accessed August 2023

(a) Overall usage *Grafana* dashboard

(b) Alarm *Grafana* dashboard

Figure 3.14: *Grafana* dashboards for data exploration.

the author.

CHAPTER 4

Results

This chapter is dedicated to presenting the results of the study, which are further grouped into three sections. Section 4.1 describes the impact of the proposed interventions on the participant's screen time and other quantitative measurements. Section 4.2 presents the analysis of physiological data, and last but not least, Section 4.3 tries to summarise emerging topics from the conducted final interviews. At the beginning of each of these three sections, a special highlighted box presents the key findings from each segment group.

Three of the original nine participants' data were removed due to incomplete and faulty usage data, leaving five male and one female student (age averaging at 23.5($\sigma = 1.70$)). To understand the impact of the proposed intervention methods for each participant, 21 days' worth of data were extracted per participant: seven days prior to the study as baseline reference (label: **baseline**), seven days as a passive observation phase (label: **passive**) and seven days with enabled nudging feature (label: **active**). For data describing participants' physical attributes (e.g., sleep), there is no baseline data available from the time prior to the start of the study. Hence, interpreting this data regarding the intervention's impact on physical data is, unfortunately, beyond the scope of this thesis.

4.1 Impact of the intervention on phone usage

4.1.1 Use of *Watchful* applications

In order to determine the degree of participant's adherence to the study scheme a look at the usage of *Watchful* application and alarm settings was taken into regard. Figures 4.1a, 4.1b, 4.1c, 4.1d, 4.1e and 4.1f have been using the phone application of *Watchful* at least throughout the study. There is quite a visible usage spike in all six figures right at the start, introduced by the necessary setup steps during the briefing. Other than that, participants have been using the app moderately during the first week

Key Findings

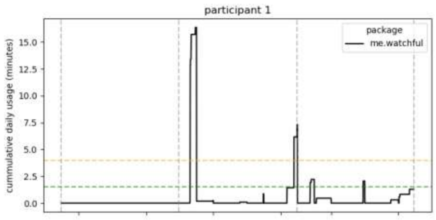
- Overall significantly reduced screen time during the first week ($p = 0.03125, e = 0.572$; see Section 4.1.2)
- Less reduction during the active week ($p = 0.15625$)
- Apps with defined limits saw similar reduction throughout both weeks (passive: -17% ($p = 0.028839, e = 0.3965$), active: -6% ($p = 0.1489$); see Section 4.1.3)
- We can see a shift when participants were using their "problematic" applications during the day:
  - A reduction between 19:00 and 07:00 for both passive and active phases (-5%, $p = 0.00049, e = 0.33$; see Figure 4.9 & Table E.7)
  - increased overall usage during the day for the active phase (between 07:00 and 19:00)
- Alarms were not as reliable as intended, and the time window for the daily limits could have been better picked, but:
  - The introduction of alarms could explain the continuous reduction in usage of limited apps between 19:00 and 07:00 o'clock; (see Sections 4.1.5 & 4.1.6).
- Phone unlock events were discarded from the analysis.

of the study. Considering that the intended primary point of interaction was the watch (*passive* phase for the first week), these results align with the expectations.

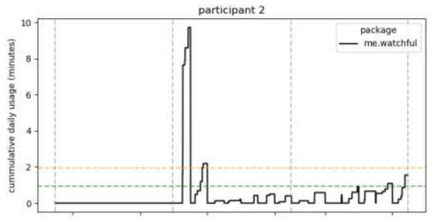
There was visible more usage during the *active* phase when participants were instructed to define alarm nudges (see Figures 4.1c to 4.1f). Participants 4 and 5 (see Figures 4.1d and 4.1e), on the other hand, did not use the app during the second phase regularly, which might be due to Participant 4 not using the nudge feature entirely. Participant 5 configured limits at the end of the first week and stuck with them until the end of the study.

Figures 4.1d to 4.1c follow the same schema.

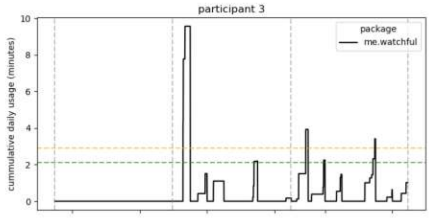
- solid line in colour: daily cumulative sum over five-minute buckets of application usage
- vertical lines in grey: to distinguish between phases



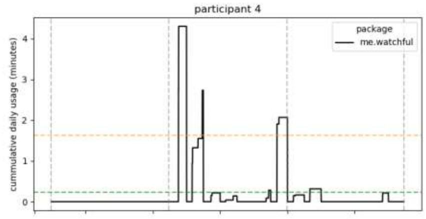
(a) Participant 1



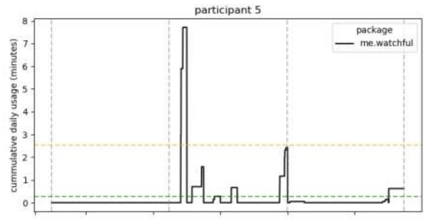
(b) Participant 2



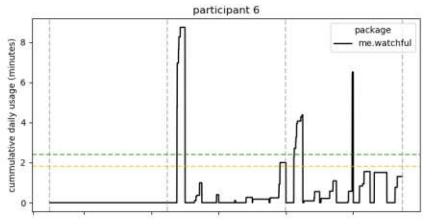
(c) Participant 3



(d) Participant 4



(e) Participant 5



(f) Participant 6

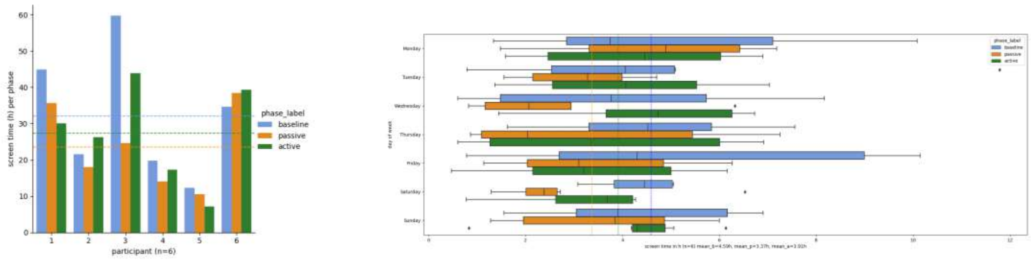
Figure 4.1: Use of *Watchful*

- dotted lines: weekly averages for phase:
  - *baseline*: blue
  - *passive*: orange
  - *active*: green

4.1.2 Analysis of weekly usage data

Figure 4.2b shows all participants' weekly overall phone usage per day of the week. We can see numerous outliers and widely spaced data points thanks to the high usage variance between participants. This can be mainly attributed to the different usage behaviours observed among participants and the relatively short study duration.

A Shapiro-Wilkes test confirmed normality ($p = 0.6083$). However, to test for significant changes between baseline and passive/active intervention phases, a modified Wilcoxon



(a) weekly usage (b) usage box-plots of all participants per day of week

Figure 4.2: Usage data group analysis

signed-rank test (a non-parametric paired test) was used due to the small sample size. Relative differences (see Equation 4.2) were calculated in favour of absolute differences (see Equation 4.1). The approach via relative differences was chosen to dampen the effect of large absolute values of highly active participants compared to changes observed with users with minor or moderate total usage.

$$\text{Absolute difference}: D_a = X_{i,1} - -X_{i,2} \tag{4.1}$$

$$\text{Relative difference}: D_r = \frac{X_{i,1} - -X_{i,2}}{X_{x_i,2}} \tag{4.2}$$

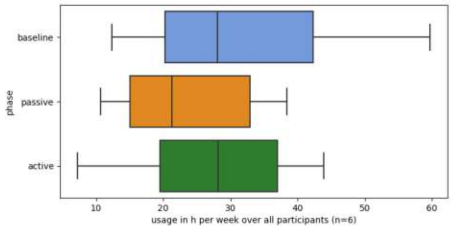
Effect sizes for significant results were calculated using *Cohen's d* estimate (see Equation 4.3).

$$d = \frac{\bar{x_1} - -\bar{x_2}}{\sqrt{(s_1^2 + s_2^2)/2}} \tag{4.3}$$

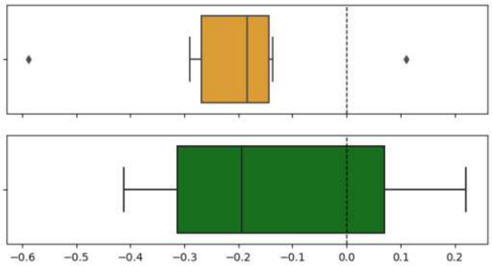
As shown in Figures 4.2a and 4.2b, daily total phone usage is highly individual and dramatically varies from participant to participant. However, we can observe a reduction between average baseline usage (4.59h) to passive (3.37h) and active (3.90h), respectively.

When comparing relative participant-wise changes between baseline and passive and active phases, respectively (see Figure 4.3b), we can observe that participants managed to reduce their overall phone usage by 21.21% on average (median 18.35%). A Wilcoxon signed-rank test on the relative changes between baseline and passive phases yields $p = 0.03125$. Thus, we can assume they are significant with middle effect size $e = 0.572$.

However, during the second week of the study, this reduction yield decreased by an average of 12.92% (median -19.436Still, due to the higher variation (see Figure 4.3b), the result is not significant ($p = 0.15625$).

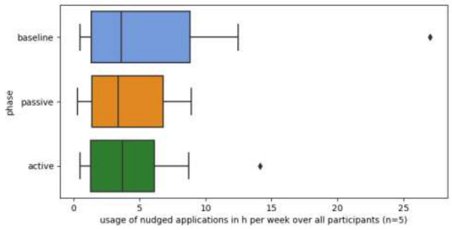


(a) Weekly time spent on phones by phases.

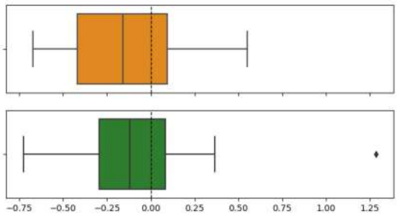


(b) Relative changes per participant between baseline and passive (*orange*,$\mathbf{p = 0.03125, e = 0.572}$) and active (*green, p=0.15625*) phases.

Figure 4.3: Weekly overall usage box-plots



(a) Weekly time spent on applications with specified limits (during active phase) by study phases.



(b) Relative changes per participant for nudged applications between baseline and passive (*orange*, **p=0.0288, e=0.3965**) and active (*green, p=0.1489*) phases.

Figure 4.4: box-plots for weekly overall usage for applications with specified limits (kept until the end of the study).

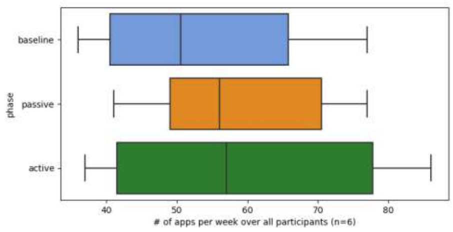
4.1.3 Analysis of the impact of alarm on the weekly usage of applications

Participants defined alarms for just a handful of applications and kept them until the end of the study. Four out of six participants configured limits for *Instagram*, *Chrome*, *Youtube*/ *Youtube Vanced* were selected by two participants each. Only one study attendee selected the application *Lovoo* as a target for alarms.

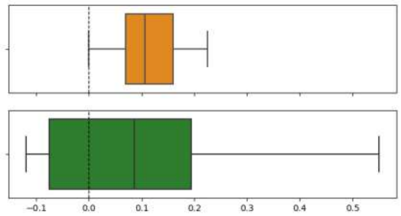
To determine the impact of the implemented nudge interventions, the weekly total usage duration of applications with a limit configured by participants during the 'active' phase of the study were selected as "nudged applications".

Figures 4.4a and 4.4b show box-plots of weekly screen time differences, respectively. There is a significant ($p = 0.0288, e = 0.396$) reduction overall between the baseline and passive phase (average -17.03%, median -15.95%).

Comparing the baseline with the active phases, changes in weekly duration were more



(a) Weekly number of apps interacted with



(b) Relative changes between baseline and passive (*orange*, $p = 0.9784$) and active (*green*, $p = 0.890$) phases.

Figure 4.5: box-plots for the weekly count of used applications.

noticeable ($average = -6.67$ %, $median = -12.00$ %). However, this change cannot be seen as significant ($p = 0.1478$), presumably thanks to the outliers.

4.1.4 Used applications per week

The number of used applications per participant per phase, in turn, increased slightly during the first and second weeks compared to the baseline (see Figure 4.5a).

4.1.5 Screen time over the day

Since screen time data was available for all three phases (each lasting seven days), examining how usage is distributed over the day per phase became interesting. Figure 4.6 shows visualisations of respective hourly screen time averages overall study participants per phase. The graphs in colour represent the respective dataset with zero values removed. In contrast, the one in grey shows the complete data to add the contribution of changes introduced by not using the device (i.e. no screen-time measurements). In comparison to this, Figure 4.7 shows daily and nightly splits, whereas *night* is defined as the period between 19:00 and 7:00 and *daily* and vice versa.

Mean hourly screen time changed (*grey*in Figure 4.6) from 10.92 minutes baseline to 8.39 minutes ($p = 0.00002, e = 0.51$) to 9.74 minutes respectively during the active phase. Whereas when looking at the cleaned set (*blue*, *orange* and *green* in Figure 4.6), more significant changes can be seen between baseline (16.51m minutes) and passive (12.93 minutes,$p = 0.00326, e = 0.27$) and active (14.96 minutes) phases respectively.

When comparing nightly and daily activities, as seen in Figure 4.7, again, for the complete dataset (*grey*) we can see a reduction from baseline (10.90 minutes) to passive (7.88m minutes, $p = 0.00073, e = 0.47$) and active (8.70 minutes,$p = 0.00073e = 0.32$) phases. Daily activities, however, changed less favourably: from baseline (10.94 minutes) to passive (8.90 minutes,$p = 0.00049, e = 0.69$) and active (back to 10.79 minutes).

The cleaned dataset behaves similarly: From 16.51 minutes (baseline) to 12.93 minutes (passive, $p = 0.00326, e = 0.27$) and 14.96 minutes (active); Nightly split: from

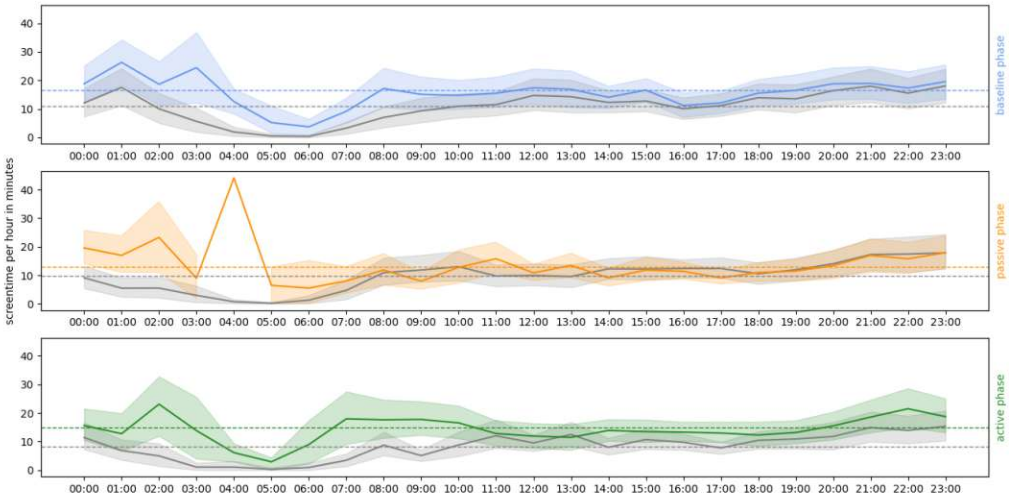


Figure 4.6: average screen time usage per hour per study phase (95% confidence interval of null-value cleaned in colour, complete dataset in grey)

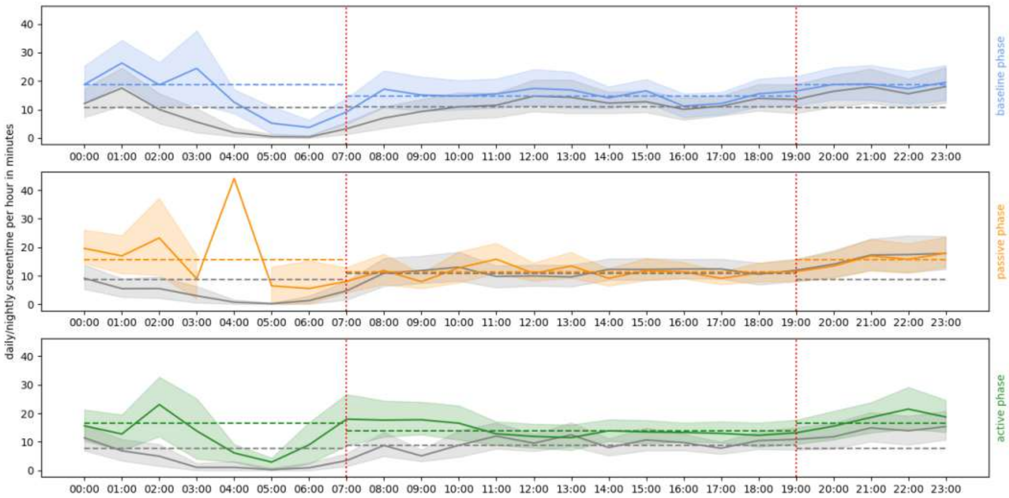


Figure 4.7: average screen time usage per hour per study phase in daily and nightly splits (95% confidence interval of null-value cleaned in colour, complete dataset in grey)

18.8 minutes (baseline) to 11.22 minutes (passive) and 16.45 minutes (active, $p = 0.00171, e = 0.39$); Daily split: from 14.72 minutes (baseline) to 15.60 minutes (passive, $p = 0.00024, e = 1.45$) and 13.94 minutes (active);

Consequently, it seems the study impacted nightly usage behaviour slightly more than during the day. See Tables E.1 and E.3 for additional details.

4.1.6 Ratio of limited application screen time to total screen time over the day

Another perspective on usage behaviour can be provided by comparing the hourly ratio of screen time caused by applications with defined limits to total hourly screen time. Looking at these relative values, rather than absolute screen time in this case, might help to make individual behaviour more comparable. Figure 4.8 shows the hourly average of screen time of limited apps to the total screen time ratio per study phase. Again, the grey graphs represent complete data; daily averages moved from 26.51% (baseline) to 23.83% (passive) and 26.17% (active), respectively. Which, in turn, matches the results seen in Section 4.1.2. In comparison to this, there are only very subtle variations in the cleaned hourly dataset: 57.35% (baseline) to 56.95% (passive) and 56.53% (active).

Analysing the daily and nightly splits, however, reveals a more significant reduction: During the night for the complete dataset: from 26.49% (baseline) to 21.84% (passive,$p = 0.00049, e = 0.33$) and 21.98% (active, $p = 0.00049, e = 0.33$). For the daily split, we can observe different behaviour: From 26.53% (baseline) to 25.83% (passive) and again increasing to 30.38% (active). This pattern is repeated in the cleaned data set: nightly split averages range from 61.77% (baseline) to 61.85% (passive) and 61.85% (active,$p = 0.00464, e = 0.52$); Daily split shows again less changes: 53.51% (baseline), 53.35% (passive) and again an increase to 56.09% (active).

Consult Tables E.7 and E.9 for more details in this matter. The significance level is very much debatable, of course, thanks to the relatively small study population.

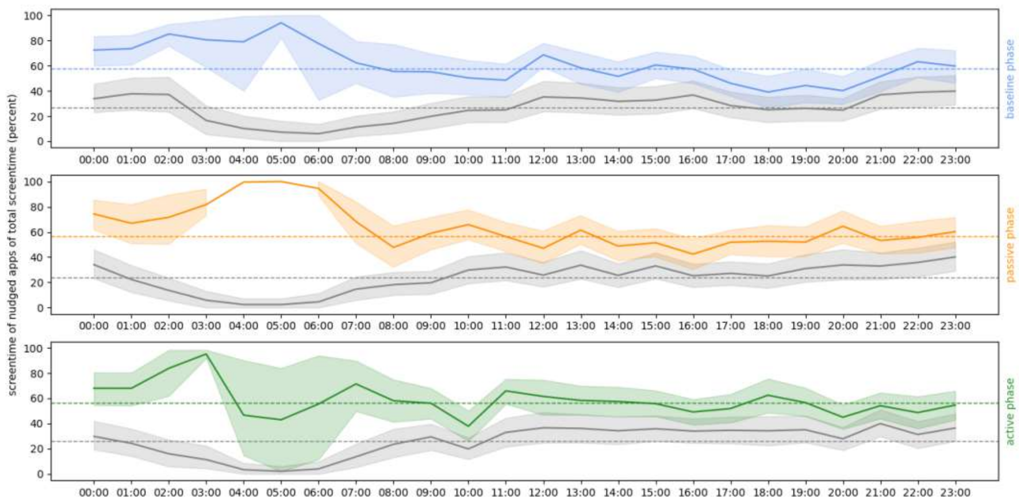


Figure 4.8: average screen time usage per hour per study phase (95% confidence interval of null-value cleaned in colour, complete dataset in grey)

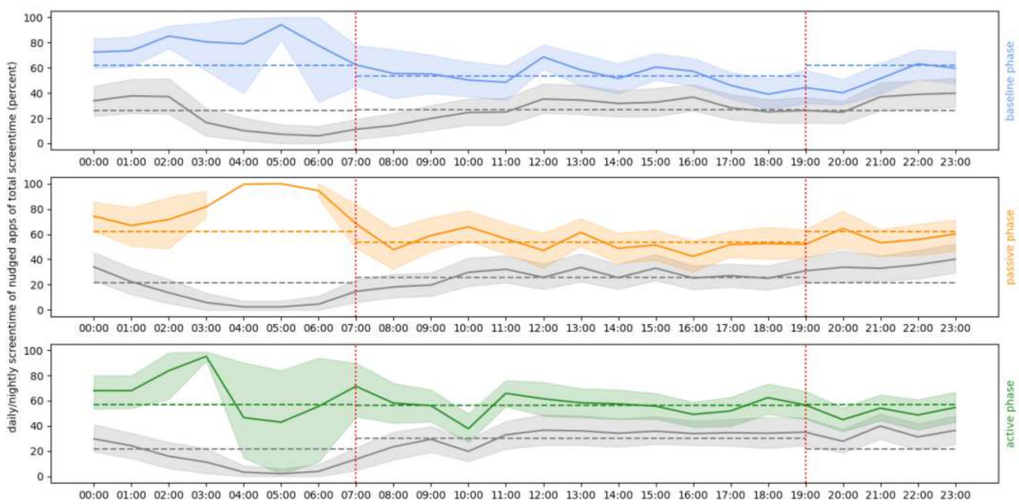


Figure 4.9: average screen time usage per hour per study phase in daily and nightly splits (95% confidence interval of null-value cleaned in colour, complete dataset in grey)

4.2 Observations on health data

Key Findings

- Physiological sensor data (heart rate & steps) was too sparse for more comprehensive analysis and have been dropped.
- The need for redundant data recording modalities becomes apparent.
- Sleep data showed consistent late bedtime behaviour throughout the study population (university students in their 20ies).
- Alarm time window could be better aligned using sleep/wake cycle data; see Section 4.2.2.
- With dense usage data or recurring usage patterns, sleep activities could be estimated.

4.2.1 Health data

Thanks to various sensors built into the watch and integration of *Google Fit*, the collection of heart rate, step and sleep-cycles data was available to be used in combination with application usage data harvested from Android's system APIs inside the *Watchful* phone application. As part of our Watchful application, health data was used for the purpose of informing the application's users about the interplay between some of their physical characteristics and digital habits. The health data integration received mixed feedback, which will be discussed in Section 4.3.2, and additionally, due to irregularly continuous data with wide stretches of missing data points, the analysis of the available health data sets will focus solely on sleep cycle/bed-time data.

4.2.2 Sleep Data

Sleep data provided by Google Fit is represented in a *session* data type, which represents a time interval of activity, optionally supplied with the sleep stage: Awake, out-of-bed, sleep, light sleep, deep sleep and REM As shown by Svenson et al. [SCT+19], wearable sleep trackers lack accuracy when detecting sleep phases and are to be viewed cautiously, but sleep trackers might still provide good results for overall sleep time. The study was done with a different watch model so some limitations might apply. Likewise, multi-modal appraches as proposed by Massar et al. [MCS+21] who use sleep tracking data from wearables, tappigraphy data from smartphones and survey based data, have shown that each modality can provide redundant and complementary information regarding sleep behavior.

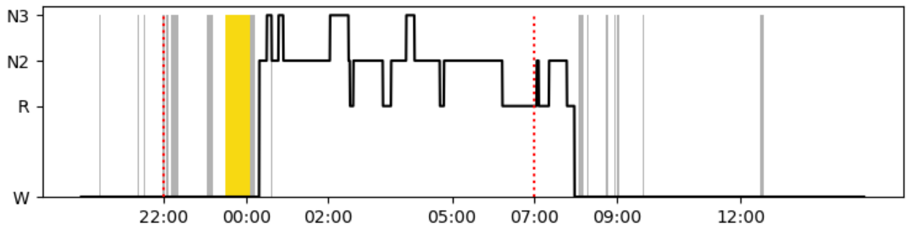


Figure 4.10: Exemplary plot of combined usage and sleep data

Sleep-tracking data will be explained with the plot visible in figure 4.10, which will display the data of one night as an example. The x-axis represents time, and y the sleep cycle (with W is Awake, R stands for REM, N_2 and N_3 for light and deep sleep, respectively). Sleep activities are represented by the black solid line, dotted red vertical lines stand for *10:00pm* and *7:00am* as reference times, and smart-phone usages with grey and app-specific colouring. In this example, the participant has been using *TikTok* right before falling asleep around midnight. The applications highlighted in the plots are those for which the participants had configured alarms for.

Figures 4.11 to 4.16 show all sleep and usage data plots for all participants for the duration of the study. It's quite obvious that sleep and phone usage have a negative correlation, which is expected. However, since 36.9 % of plots are missing sleep data, only speculations are possible. Still, rough estimations of sleep activities can be easily done by looking for large enough gaps between usage data points, depending on the density of usage data points.

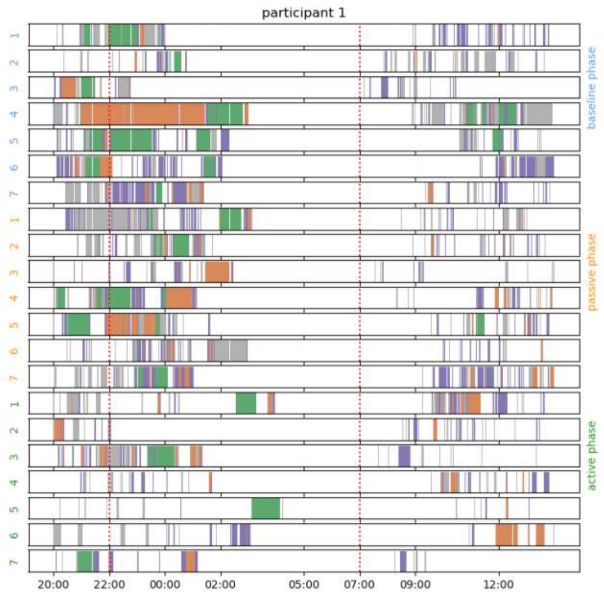


Figure 4.11: Participant 1' sleep: *Instagram*: violet; *Chrome*: orange; *Youtube*: green

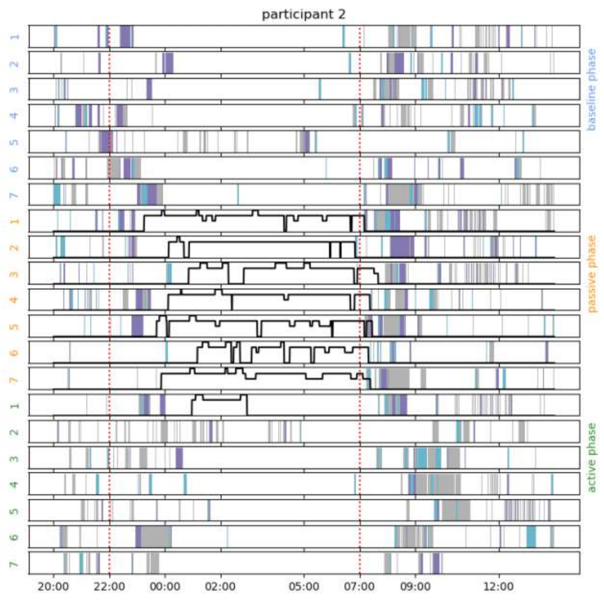


Figure 4.12: Participant 2' sleep: *Instagram*: violet; *Lovoo*: turquoise

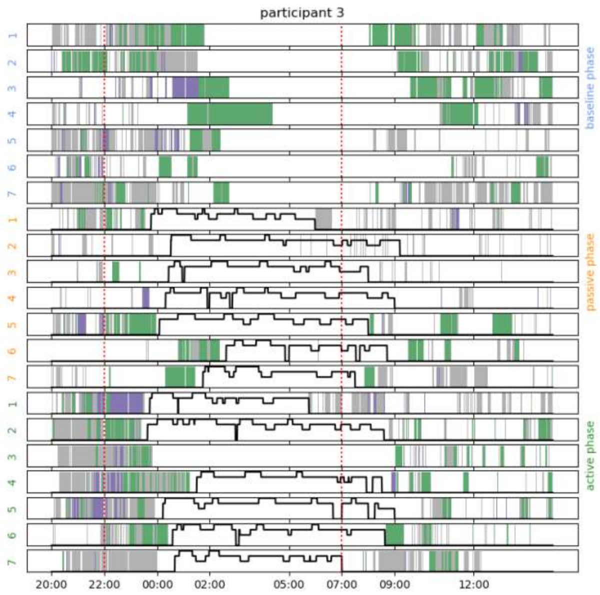


Figure 4.13: Participant 3' sleep: *Instagram*: violet; *Youtube*: green

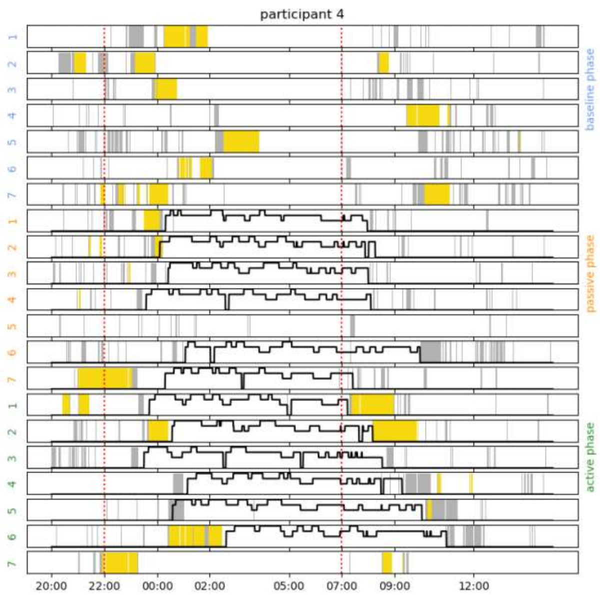


Figure 4.14: Participant 4' sleep: *TikTok*: yellow

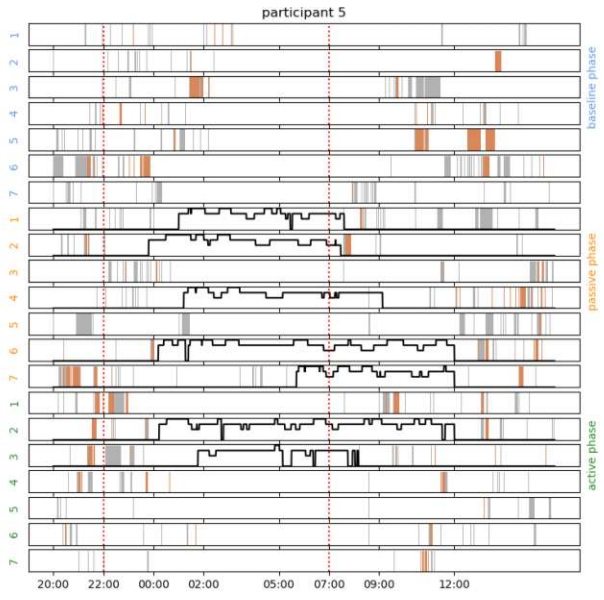


Figure 4.15: Participant 5' sleep: *Chrome*: orange

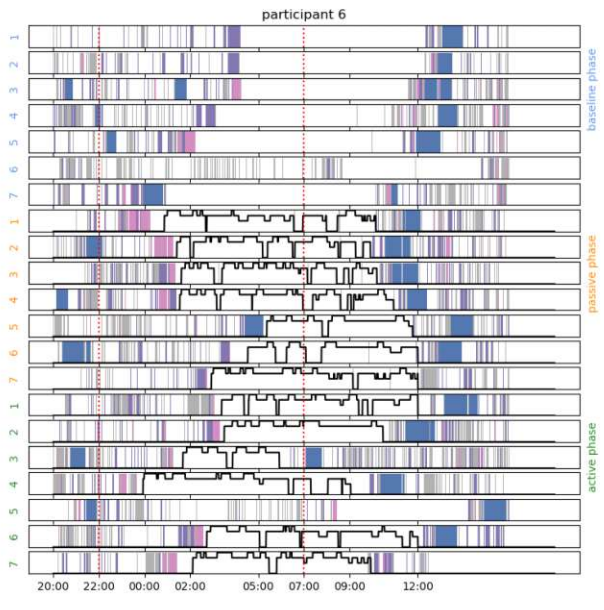


Figure 4.16: Participant 6' sleep: *Instagram*: violet;*Amazon Prime*:blue;*Youtube*[1]:pink

4.3 Observations from survey and interview data

Key Findings

- Participants liked the watch interface, but wearing the watch felt obstructive to some; see Section 4.3.3.
- Participants voiced positive overall feedback despite being annoyed by alarm vibration; see Section 4.3.6.
- Participants claimed the watch might have reduced interaction with their phones by reading notifications on the watch instead; see Section 4.3.4.
- Participants were not convinced about having health & usage data combined but liked the watch's sleep-tracking features (see Sections 4.3.7 & 4.3.8)
- Alarms should be further improved in terms of transparency and accountability; see Section 4.3.9.
- Participants felt left alone in the process and wished for more guidance and appraisal; see Section 4.3.9.

4.3.1 Survey

Participants were asked to fill out the survey at the end of the first week of the study to capture initial experiences and address technical issues in the early stage of the study.

See Tables A.1, A.2 and A.3 in the appendix for more information about the nature of the questions.

However, due to the low number of participants ($n = 6$), no statistical significance can be expressed with the collected responses (see Figures A.1a to A.1i in the appendix).

For some of the questions, trends are imaginable towards one or the other end of the spectrum of experiences. However, with better qualitative data made available from the recorded interviews, the feedback and opinions presented in the next section are based on those.

4.3.2 User interviews

The interviews followed a semi-structured approach to encourage participants to reveal their personal experiences during the two weeks of the study. Although the prepared

[1] Youtube Vanced

interview guideline (see Section B in the appendix) focuses on the prototype's available features, participants added their own ideas and hypotheses from their ventures.

Thematic analysis was used to grasp common patterns and themes from the available body of recordings. A Miro [2] dashboard was used by me to collect and cluster common themes, which will be presented in the next section.

4.3.3 Wearing the watch

Participants described how wearing the watch reminded them of why they got it in the first place (the study's general topic of digital detox). Two coping strategies with this reminder were identified: first, we can observe a degree of habituation and falling back into old usage patterns, as this participant describes it: *[. . .] Maybe just like being shown how much you use your phone on a daily basis, which the watch facilitates– you're consciously thinking about it. But after a while, it just gets thrown into your subconscious again. So you don't really think about it [=your usage] [any longer].*

The second strategy is adherence, as described by one other participant, who was bothered so much by the watch throughout the study that it stuck with him for a longer period: *I used my phone a lot more before that.[. . .] I wasn't used to wearing a watch. And so it was kind of like you have something that always reminds you why you're wearing this [watch] [. . .] I don't really like the watch, as my wrists are too small for the watch. That's the first one. And the second one is it's not really my kind of style.*

Wearing the watch throughout the day and night was the one thing participants mostly complained about. One participant couldn't sleep while wearing the watch, and others felt uncomfortable or experienced the watch had run out of power in the morning: *Sometimes it's a bit uncomfortable to wear a watch during sleep*

4.3.4 "Handsfree notification access"

Modern phones can be used for many tasks, from communication to taking pictures or consuming various content. A wristwatch, on the other hand, serves the most basic purpose of providing easy access to information about the time of day. Something that mobile phones have been used for more prevalent in recent years — one participant reported that out of habit, they used their phone to check the time, despite wearing the watch. One participant described having the watch stopped them from getting side-tracked with other things on their phone when handling notifications there: *It helps to get notifications on the watch to determine their importance without actually touching the phone. Because when I lay my hand on my phone to check a notification, it's never just checking the notification. I spend five, ten minutes there — these little detours sum up over the day. [translated analogously from German].*

Another participant described their experience in the following way: *I would say that people tend to unlock their phone less if they have a watch just because you get notifications*

[2] Miro, `https://miro.com/`–Accessed January 2024

and you can check the time using the watch. Like I think most people they use, they're — even if they just want to check the time or they're bored, they will look at their phone and then say, Oh, I got a new notification. Then they unlock it. So, especially with certain apps that like get push notifications, they give push notifications, so you do engage. So yeah, I think the watch can curb that if that's what you want to do.

Almost all recent smartwatches offer features to relay notifications from the phone to wearers. Since notifications tend to trigger interaction with the phone as they either signal users they need to react to something urgent or they carry information, which can develop some highly annoying character, as one of the interviewed participants put it:

I limited my notifications to the watch just because it gets annoying, especially if you're in huge group chats or something. It vibrates all the time, which is really off-putting. [...] But you can customise it [=the notifications]. So that's fine Or: *Sometimes, when I look at my phone because of a notification, I end up scrolling through Instagram or reading emails. That's so unnecessary.[translated analogously from German]*

4.3.5 Raising usage awareness through monitoring

Some participants reported being surprised at how much time they spent on the phone, with certain apps, or how many times they unlocked their phones over the day: *And what was more a statistic that made me see how many times I used to unlock my phone, which I didn't expect it at all. It was like 100 times or more. And that's a problem at least in my eyes. And yeah, I think the app does a good job in displaying all the statistics really well. It's good to see, at the end of the day, how much time I've spent on apps. Honestly, I was surprised at how many hours go into certain applications.[translated analogously from German]*

Other participants either already had a decent understanding of their usage or weren't that much concerned about it. However, they speculated about how making users more aware of their screen time might help them to reduce it.

I didn't really use the app all that much, but I thought the watch face was pretty useful because even just taking a glance to look at the time, you see how much screen time you had on the day. And you can try and correct a little bit if you figured it was too much. So I already knew my pattern, kind of. And that's why I also don't check it regularly. But I think it can be helpful for people that use the phone more often than me and also excessively to see when they use the phone during the day and when they actually can minimize it or even [turn off] the phone. Because it's just an accessory.

One of the reasons why participants tried to stick to their motivation towards limiting their phone or app time was described as they felt being judged by monitoring applications and didn't want to put up a negative record in any way: *it's because you see how much you use your phone and then this generates a bit of guilt. You don't tend to use it that much anymore. I think in my case at least.* Or, as another participant put it: *It's just like it's watching me [...] I don't want to put up a high score or something. So*

I just use my phone extremely little. I don't know wh . . . Normally I use it for 2 to 4 hours, something like that. And it's just. I don't know. I think it was like kind of a new environment when wearing this watch. And in combination with the application and the alarm, which was annoying — to say it in kind words — This really helped to lower my phone usage.

However, although there is some negative smack regarding them being tracked, participants were able to take some useful insights away from the experiment and even left positive notes:

It's always nice because you can keep track. It's like working out, and you have a journal you can always see.

I used to look at my phone usage, but not every app — analyzing each app — how much I was using it. But with these two weeks, I think it got more and more clear to me.

I think this overall monitoring already reduced my phone usage in the first week because I was more aware of my usage patterns. And this — so the limitation that I could set for myself reduced it further.

4.3.6 Alarms

The usage limit nudges received equally the most love and hate from participants. Even though alarms weren't triggered as often as they probably could have if configured properly, there is general acceptance of the fact that alarms need to be annoying and disturbing in order to work: *I think they get the job done because they're extremely annoying, which is probably their purpose.*

Another participant described how the recurring alarms after reaching the limit triggered a feeling of having done something bad, which stopped further usage: *The feature I liked the most and least is the same: The alarm. It was . . . The first time I heard it, it was so annoying I wanted to throw the watch away.[. . .] When I opened chrome again after that, the alarm went off again.[. . .] That deterred me because I don't want an alarm going off every time. Like, I don't know, I feel like I've escaped jail or something else.* Others voiced similar experiences: *Most useful was the vibration on the watch when I overused too much. Because it was really, you know, it made me kind of angry with the watch. But that was the point. [. . .] every time it vibrated, I didn't want it to repeat, so I didn't even use the app at all. So I just forgot about it. Yeah. So I think that's a good sign. [The alarms] really helped me, because I think it's not a good habit to use Instagram for two or three hours per day, for example. By setting an alarm I get reminded that it's enough for the day. [translated analogously from German]*

It's quite useful but sometimes annoying. Yeah. [. . .] I think it's a great idea.

Ultimately, the nudge's usefulness seems to outweigh the level of annoyance that the feature presents in order to reduce bad habits. The most noteworthy limitation of the alarm feature is that due to the short study duration, only short-time effects can be observed.

4.3.7 Simple interfaces and unobstructed user experience

Participants found positive words for the prototype's user interface and the general design of the companion watch app. Providing quick and unobstructed access to information through the watch and, if necessary, further details by accessing the phone application following consistent user interface design was received with kind feedback.

I didn't really use the app all that much, but I thought the watch face was pretty useful because even just taking a glance to look at the time, you see how much screen time you had on the day. And you can try and correct a little bit if you figured it was too much.

So this [the watchface] was quite nice just because it was more convenient than to open it on the phone and look it up. Uh, the only thing I didn't really use in the app was aggregating the data for weekly a overview.

I think it is pretty clear, especially with the colours and the dial. it gives you a sense of all of you like, okay, I'm using this much of this app, and I should be using it less.

The overview is ok because I can see which application has which colour. On the watch face, I sometimes forget the colour scheme of my most used apps. Otherwise, it's good. [translated analogously from German]

It's an easy way to look [at] the usage instead of going to the [system] settings I think that when I wanted, like, more detailed data, then I use the phone app, but when I just wanted to look at my 'minutes', then I would actually rather use the watch because it's closer in reach — it's easier to read a little bit.

4.3.8 Raising health awareness

Health-related features provided through the watch's preinstalled software or *Watchful* received mixed feedback. This ranges from disinterest to high motivation to watching health signals change over time or through activities.

For example, this participant saw no usefulness in the health sensor data:

Sometimes I would go: Oh, that's my heart rate? Cool [ironic tone].

Others described keen interest and even interpreted the data to improve their well-being further.

I like, for example, that I could see how my heart rate changed during the usage of an app. So I could see the excitement that caused the item — the appeal of the app to use it, for example. And I also like that I could see my overall health a bit better. So I like to know when I'm stressed or how I could sleep, for example. [...] I think that I've improved my sleep a bit using this app or the watch overall.

I think — the most used features where the heart rate sensor and the sleep tracker [on the watch trough the Suunto app]. Because it was like, okay. And now we have a goal to live up to.

Some criticism about the provided sleep score (calculated by *Suunto*) quality was voiced by participants, which was also briefly discussed in a different context in Section 4.2.1.

So heart rate and sleep data and so on were quite interesting. Especially the sleep data — it always said that I had a bad sleep, which I do not always agree [with]. But it's kind of good to see how my sleep rhythm is visualised.

On a side note, participants refer to both health apps available on the watch (*Suunto* and *Google Fit*) and the *Watchful* app on the phone. Which makes the distinction altogether impossible to which degree the prototype app implementation has impacted health awareness.

4.3.9 What's missing according to participants?

Participants described some additions to existing features or new ones that they felt would improve the user experience or further help reduce screen time.

Making the internal state of limits more transparent

The first group of issues describes participants' need to understand better how far away they are from reaching their set limits. Some propose using notifications for this purpose; others would also like to see this integrated into the user interface (the watch face and the phone app).

Like, if you set up a threshold for an app. I don't know if you could, like, maybe colour code how far away you are from that threshold.

Like, I want to use the phone for 3 hours a day, and that's it and maybe you just get every hour [a] notification: You've used it for one hour — for 2 hours — as it gets closer to the 3 hours [limit].

Maybe reminders or something like that, that if you set the alarm to be half an hour or so. That you can set, also reminders to give you a notification that you have already spent — I don't know — 0 minutes using that.

Both feature recommendations are certainly justified and should be incorporated into future development activities, given that an appropriate design can be found for the watch face implementation.

For the participants less convinced about interacting and using a watch, one participant described integrating a widget into the phone, providing similarly easy access to information about daily screen time.

Maybe you could, um, when you like, when you have an app open and if you, like, pull down the hot bar, you could have, like, a permanent notification. So where it says, like, okay, you have spent 20 minutes in this app. If you wanted to check while being in the app.

The described functionality matches features already implemented in *Digital Wellbeing* and *MyTime* [HHKK16] and provide some useful addition or fallback to the prototype implementation.

Snoozing alarms and using 'incognito mode'

The only way for participants to continue using applications with configured limits after reaching the maximum allowed threshold (90 minutes) is to delete the limit. Otherwise, the alarm will be triggered on the watch continuously every 10 to 15 seconds. Which had the prior not anticipated effect that limits were not being re-created in a timely manner afterwards. In some cases, it took a few days. See Figures 5.2 for an example of this described behaviour.

One participant describes the strategy to deal with the situation of prolonging the alarm limits in their case, while others have resolved to delete the alarm entirely instead, as we can see from the phone app metrics.

But there was even times, like when I needed to use the app, so I just made the [allowed] usage a bit longer. [. . .] Maybe like a 10-minute extra time for an app, and that's it.

Another participant described the idea of being able to add a timeout or a timed exception function, which would solve the need to change the limit's value to keep using an application for a little more time.

Maybe there could be a way to dismiss the next 15 minutes of YouTube from counting towards the alarm limit because I'm just listening to music while showering. [translated analogously from German]

This and another comment by the same participant highlight how possible users might feel misunderstood by the alarm logic and how it assesses screen time. *There are situations where the alarms are unnecessary; for example: for cooking, I'm using Google to look up recipes. From those 15 tracked minutes, I really used the phone for one minute, and then the alarm got off nonetheless. [translated analogously from German]*

The participant expresses the need to understand the justification for a raised alarm, which was apparently triggered by one of the participant's 'wrongdoings'. If the participant finds the ruling unjustified or simply wrong, adherence to and acceptance of the intervention might suffer in the long run.

Guidance and appraisal

Participants voiced the desire to have more guidance on reducing screen time. One suggestion was to recommend pre-defined alarms for the most used applications over the last week.

You could do something like where you see the — where you put in pre-configured alarms for the most used apps of that person. So the first week, you analyse usage using the *watch, and the second [week], you say, "Okay, these are your top three apps. Would you like to set up an alarm for these apps"?*

This approach could be reused in consecutive weeks to build iterative adaptations to the alarm set, further limiting the risk of defining alarm limits that do not represent the recently observed usage behaviour.

One other suggestion was to recommend alternatives to spending time on the phone or in apps via notification or alarm text on the watch:

Maybe if you are active for a long time on the phone, you get a notification or reminder that you should maybe turn off the phone and take a walk or something like that. Yeah, that would be nice.

One final appeal left by a participant was that they felt long-term goals and appraisal for doing better over time were missing and might add some motivation to reduce screen time further.

Something like an app-specific graph. Like, if you want to use, for example, Facebook a bit less than it would be useful if the app had some summary of only the usage of, for example, Facebook. And then you could set some goals for the week or so maybe or get some notifications that you did better this week and use it less or something like that.

CHAPTER 5

Discussion

Where Chapter 4 presents an overview and highlights the results from the user study, this upcoming chapter tries to discuss these results in light of comparable studies found in the literature (see Section 2.4) and the proposed research questions from Section 1.3. The chapter starts with discussing factors that might have supported the goal of reducing phone screen time, followed by factors and observations that could have mitigated the same.

5.1 Factors supporting a reduction of screen time

Reducing phone and specific application screen time was successful for some participants. Throughout the study, several factors that have contributed in a short-term manner to the observed results are collected in this section:

5.1.1 Increased awareness

Participants indicated that participating in the study raised their sensitivity towards when or how much time they spent with specific applications. The prototype application and the watch provided simple and unobstructed access to usage statistics, allowing participants to make informed decisions by monitoring their smartphone usage.

5.1.2 Novelty effect of the watch

Wearing a smartwatch or watch was unaccustomed to some participants and thus served as a reminder of why they had been provided with the device in the first place.

Although the feeling usually declined over time and was expected to occur, in some cases, the effect on usage reduction might have been stronger than anticipated, especially for participants who felt obstructed by the size and weight of the watch.

5.1.3 Setting realistic limits and study schema adherence

We have seen examples of nudged applications that were configured with limits close to or below past weeks' average daily usage (for instance, see Figures D.1a, D.1b and D.3a). Here, the alarm successfully mitigated the use of specific applications beyond the intended limit, thus supposedly helping to reduce the average daily screen time for selected applications.

Due to the similarity of the alarm design to Okeke et al. [OSDE18], it can be expected that the alarm in the proposed design does not support the building of sustainable long-term habits. It is possible that, when observed for a longer duration, the forceful alarm intervention could lead to abandonment, just like Okeke et al. [OSDE18].

5.1.4 Fear of consequences

In multiple cases, participants almost unanimously described alarms as incredibly annoying and too forceful. Besides respecting alarms and the reached self-set limit they represent, one participant described the first occurrence as so perturbing that avoiding the alarm signals entirely became a goal for the remaining part of the study (see Figure D.5).

Admittedly, this was not the intended effect of the alarm nudges. Yet, the observed behaviour is still insightful and should make developers and researchers even more considerate about nudge design — see "nudge for good" [Tha18]. Guidelines and recommendations found in the literature and published reference implementations (see Sections 2.4 and 2.8) could have mitigated this design flaw for this prototype to some degree.

This effect is likely to be short-term and probably lead to abandonment, similar to what Okeke et al. reported about the long-term sustainability of their proposed intervention [OSDE18].

5.1.5 Inherent smartwatch features

The watch and its available features might have reduced the need to interact with the phone occasionally. With a watch as a provider of time information and passed-through notifications, using the smartwatch might reduce the risk of 'wandering off' since occasions to interact with a phone become less necessary. Oulasvirta et al. reported that phone-checking habits could be a gateway to other applications and consequently extended screen-time [ORMR12].

Yet the attention cost of receiving a notification on the watch might still be similar to one popping up on the phone [SMY15, MPV$^+$16, MBL$^+$17], so careful consideration should go into selecting applications capable of pushing notifications to the watch.

5.1.6 External factors

External factors or situations beyond the user's control might influence screen time reduction beneficially, such as having less free time and being in social company. Examples are periods of increased workloads for university and work, simply time spent with socialising or sporting activities in the real world. Travel might also offer an occasion to evade everyday usage patterns or move to different times during the day.

Since all participants are computer science students and parts of the study fell into the last week of the semester, traditionally filled with examinations, tests and end-of-semester festivities, we can expect phone usage behaviour to be impacted by this to an unknown degree. Although none of the participants reported anything "out of the ordinary" during the debriefing interview, they might have deemed factors not mention-worthy or slipped the participants' minds entirely.

5.2 Factors mitigating a reduction in screen time

Like the last section, this one summarises observed factors that might have mitigated the successful reduction of screen time on phones or specific applications.

Effects and motivation wear off

Undoing learned habits takes time and patience, three months and upwards, to become sustainable [LVJPW10]. Facing slips into old behaviour and reflecting is part of this journey, as it is hard to maintain motivation and discipline over time, especially when practising moderation and self-regulation.

Consequently, one can only expect a short-term effect from two weeks of interventions. If any, at all. Relapses and fluctuations in application usage are widespread, especially when starting with changes in phone habits. On average, the reduction in overall usage was more significant during the first week of the study (see Figures 4.3a and 4.3b) in comparison to the second week despite the introduction of a new intervention, which could be already interpreted as the commitment to stick to detox undertakings is wearing thin. As we have seen with Participant 2, who disabled Bluetooth during the second week of the study, effectively killing alarm and watch-face data updates, usage has increased again.

Another aspect of this study is that wearing our prototype watch can be a nuisance, as some participants mentioned in their final interview. Thus, the watch plays the role of a constant reminder. However, this novelty effect also wears off with time while getting accustomed to the watch. Consequently, a relapse into no longer desired behaviour becomes probably more likely. Other studies have found similar doubts on the long-term sustainability of their proposed digital-detox interventions [HHKK16, OSDE18, KJKL19, MRDR19].

Furthermore, as most recently described by Parry et al. investigating the use of digital detox: “A majority of those who do use digital well-being applications can be described as passive-occasional users for whom objective information about their phone use patterns is exciting but does not serve to initiate efforts to change behaviour” [PlM+23].

Adherence to the study schema

During the study’s second phase, participants were instructed to define screen time limits for applications they deemed troublesome. One participant did not define any limits for the active phase, whereas others had restricted their application time already during the passive stage. In the first case, we observed a reduced overall phone screen time, but there was little change in the most used applications. The latter case could have skewed results in favour of the passive week as impacts on reduced usage have been attributed to the passive factors.

With the current implementation of the prototype, disabling alarm functionality until the beginning of the second week was not intended. In retrospect, however, this would have avoided the observed issues.

Choice architecture and alarm design

By definition, nudges should offer options without prohibiting choices [TS08]. The implemented alarm nudge offers the choice to stop using an app, dismiss it, or ignore it entirely. In the two later cases, the participant’s choice will get ‘punished’ either through a newly triggered alarm shortly after or a never-ending ongoing alarm. Participants called that rightfully too forceful and questioned the design.

Ultimately, negative reinforcement has not been recommended to support behaviour change and healthy habit formation [APS97, PBS+23]. We can see the effect indirectly by participants moving to the uppermost available limit setting to evade the alarm.

User-defined alarm limits

Some participants defined limits that were above their average usage and would only serve as an ultimate keeper to mitigate binged usage, thus being ineffective in lowering screen time in the long run. Limits could have been set to more fitting values automatically, taking past weeks’ average usage into account as a reference to take effect in reducing usage. Otherwise, the detox application should support the participant in becoming more aware of limits that meet the participant’s past usage to make change easier. However, we can assume that the first week of the study did not raise usage awareness in a way that motivated participants to determine limits that represented their daily application usage, which aligns with the observations of Parry [PlM+23]. The participants’ personal commitment to undergo digital detox was not part of the selection criteria during participant recruitment. The sole interest in digital detoxification interventions and the ownership of an Android device made it easy to approach and recruit from a student pool. However, as already discussed by Parry, for the majority of users, information on phone use patterns is interesting but does not serve to initiate efforts to change behaviour [PlM+23].

Additionally, we can see that in some cases, alarm settings got altered and increased when limits caused too much annoyance and, in turn, were rarely decreased again afterwards. Letting participants define their own limits and allow them to edit limits at any given time might have diminished the efficacy of the alarm feature itself. Still, it provides some insights into managing self-targeted restrictions. Other studies, which handled this regard differently by setting a hard limit based on past usage [OSDE18, SMWM22], might have made it easier to establish the efficacy of their intervention methods. However, inconvenient or too strict restrictions might lead to intervention abandonment despite having the best for the participant in mind [KJKL19, HHKK16].

Lyngs et al. found in their review that the expectancy component (how likely the user thinks she will reach their goal) was less frequently implemented and was highlighted as an under-explored area [LLS+19]. Villalobos et al. evaluated digital change behaviour applications in the eyes of self-determination theory and found that only 25% of the sample provided support for all the basic psychological needs. Features like reminders and self-defined goal settings support a user's basic need for autonomy, whereas externally given goals could undermine that [VZC20].

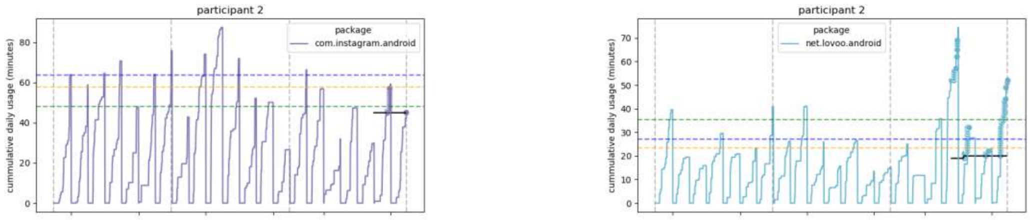
Finding the balance between self-defined and suggested goals could be solved through user-tailored interventions.

"Daily" time window assumptions

The available sleep data (see Section 4.2.1, Figures 4.14 to 4.13) suggest that the majority of participants indulge in bedtimes well after midnight. With the current prototype's implemented definition of the *daily* time window, spanning from *00:00:00* to *23.59:59*, this might have introduced edge cases regarding app-usage calculations. Resetting the daily total app-screen time to zero at midnight allowed continuous use of applications without triggering the alarm.

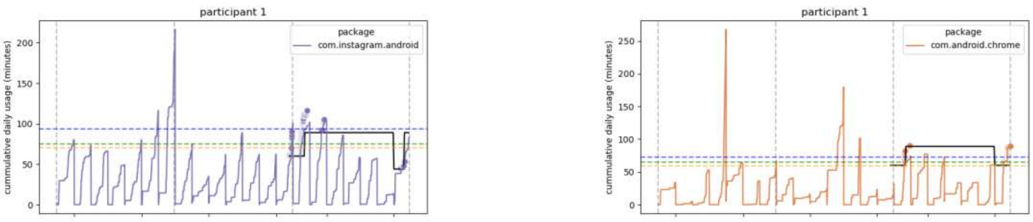
This assumption might not have reduced the total effectiveness of the alarm nudges since the usage was still tracked, but it might have mitigated alarms when they would have been valuable and well-timed. In a way, this aligns with the findings of Kim et al. [KJKL19] that users would wait for the lockout to reset at midnight and continue using their apps.

Especially when it is considered essential to limit screen usage in the hours before going to sleep [MHZ+19] (see Section 3.6 for the average bedtime) The usage of the applications that participants considered alarm-worthy generally increased between 22:00 and 01:00 (see Section 4.1.5). In retrospect, resetting the limit cap during that critical period before bedtime seems unwise.



(a) *com.instagram.android/Instagram* (b) *net.lovoo.android/Lovoo*

Figure 5.1: Participant 2's usage of *com.instagram.android* and *net.lovoo.android*



(a) *com.instagram.android/Instagram* (b) *com.android.chrome/Google Chrome*

Figure 5.2: Participant 1's usage of *com.instagram.android* and *com.android.chrome*

This assumption and the observed behaviour highlight the requirement of a similar understanding of what is "*daily*" between the application and the user target group.

Technical issues

Technical prototypes seldom work 100 percent as expected when released into the wild, and this prototype was no exception. One participant described that synchronisation with the watch did not work out of the box after the phone got restarted until the Watchful application was started, which points to implementation issues with the background worker on the app's side. With the worker jobs disabled, the usage data displayed on the watch face would become stale, and the evaluation logic for alarms would not execute, rendering the prototype practically useless.

Further issues arose when one participant disabled Bluetooth to preserve the battery, cutting the communication between the phone and watch and consequently rendering triggered alarms pointless (see Figures 5.1a and 5.1b and in comparison Figures 5.2a and 5.2b). The alarm signals got dispatched but never reached the watch — apparently, the fallback via WIFI did not take effect in this situation — resulting in silent alarms that were not perceivable by the participant.

External factors

It has become apparent that some applications are more complex to not use than others. For example, apps with the capability to provide the means for communication or social interactions.

The time spent with others, the social exchange or the prospect of relations might drive the need to use the app. Furthermore, a dependency on communication or interaction partners makes planning for reduction difficult due to unforeseeable actions by these partners. Examples of this behaviour can be seen in Participant 2's data with the app *Lovoo* (see especially Figure D.2b). Despite setting a limit (which admittedly failed), the app's usage increased during the last week, which might be due to unexpected activity on the app. Similarly, we see the same pattern for Participant 3's usage of *Telegram* (see Table C.3) — there might have been increased social activity in this specific week. However, without data to confirm this assumption, that's just speculation. Still, this observation might support the point that some applications are more straightforward to detox from than others without knowing the context or what the user expects to gain from using them.

Just as some situations help break everyday usage patterns, others might allow extended phone or app usage. Although none of the participants reported anything unusual, the fact that parts of the study took place during the semester break, with all participants being university students, could indicate more spare time available than during the semester.

5.3 Summary and recommendations

Starting with answering question research question **RQ 1.1**, the alarms could not sustain the reduction level of the first week due to multiple reasons. However, despite being probably too forceful, participants responded well to the feature and claimed it helped them to reduce time on specific applications. We can see a continuing trend of reduced usage in the cases of some participants. Still, due to the small number of participants and with data for a short-lived time frame, it is impossible to generalise outcomes or assess a long-term effect. Implementing a different intervention design using positive reinforcements and better nudge and choice architecture could encourage and motivate participants to alter their screen time behaviour.

Having the watch face (question **RQ 1.2**) to offer 'real-time' feedback on phone usage was received well, as it was easier to take a look at the watch instead of opening specific apps or settings on the phone. A significant reduction in screen time during the first week of the study could be observed. Some participants reported that wearing a watch as part of the study was a physical reminder to 'reduce phone usage', thus possibly skewing some of the results. However, this novelty effect became smaller over time as participants became more accustomed to the watch.

One participant reported that using the smartwatch's notification relay feature helped reduce phone interactions. Without monitoring notifications, the current prototype does lack the means to verify this claim. Pizza et al. have discussed the watch's role as an intermediary and the possibilities to reduce time on the phone, although they might be disruptive if configured wrong [PBML16].

Consequently, to answer the overarching question **RQ 1**, there might be potential in using smartwatches to support digital detox activities to reduce screen time. However, only a few significant short-term effects on screen time with the implemented intervention and study design (see Section 3.2) were observable. To better understand possible intervention designs, a literature review was conducted to retrieve available design guidelines (as presented in Section 2.8), and an improved design proposal taking the guidelines into account has been addressed in Section 6.4.1.

The prototype presented in Sections 3.5 and 3.6 is certainly one possible design solution. The intervention design shares strong similarities with Okeke et.al. [OSDE18] with analogues outcomes regarding perceived usefulness and annoyance with the accompanying effects.

Regarding health data, the limited results (see Section 4.2.1) only allow comments on sleep data. There is a strong negative correlation between phone usage and sleeping hours for obvious reasons. However, given the fact that a large body of literature considers phone usage before bedtime as not a healthy habit for mental well-being [MHZ+19], and the findings that usage of applications deemed worthy of an alarm was increased during the night-time, there might be possibilities to design interventions using this data.

Unfortunately, heart rate and step data have been too incomplete to be used for further analysis. Henceforth, answering question **RQ 2** to a satisfactory extent is impossible with the current data and results. For future implementations, an approach using redundant ways of recording health metrics would be beneficial and help avoid the pitfalls of missing data.

Speaking of intervention designs, to answer question **RQ 3**, reflecting upon the results and observed issues with the proposed prototype and intervention design is necessary. Participants made very useful recommendations regarding what they are missing in terms of features (alarm snoozing, more transparency, and additional guidance — see Section 4.3.9).

The user experiences could be improved by fixing the observed technical issues and a revised intervention architecture ("nudging for good" [TS08]). Implementing mechanics found in the rich literature on goal setting and goal fulfilment could further enhance motivation for the right participants. (i.e. those that are inherently motivated and interested [PlM+23])

5.3.1 Recommendations

The results (see Section 4) and the preceding discussion sections of this thesis suggest that some recommendations for future designs can be added to the available guidelines that have been presented in Section 2.8.

Guidance Although phone users know digital detox as a concept, many are reluctant to try it or don't see the need to change usage-behaviour [PlM+23].

Providing users with a toolkit shows good intent, though. Still, users can achieve more by being taught how to use it properly within their capabilities, thus helping manage expectations and define sustainable long-term goals. This also means that guidance should be an adaptive process that accompanies the potential user in finding mechanisms or configurations that match their needs and contribute to their goal.

We have seen that limits were consistently set with more freedom, moved or deleted to allow for more consumption, possibly due to alarms being too forceful. However, a guided mechanism might support users in finding suitable configurations independently. This process might even adapt configuration automatically to help individuals achieve sustainable long-term goals. Research dealing with goal-setting could help shed light on providing proper support for those in need. Morrison et al. already suggested that the acceptability of self-assessment or monitoring components may be optimised by also providing tailored feedback [MMMMY14].

Redundancy The prototype design that has been proposed as part of this work is a distributed system on a relatively small scale. Despite the few components, communication between is crucial to work reliably, and communication loss to partners needs to be detected and addressed.

Intervention mechanisms should, if possible, work on either ecosystem device in case the primary one is not within reach or in a functional state.

Redundant data-gathering mechanisms could further prove beneficial to make interventions effective and efficient. Massar et al. evaluated different modalities to record sleep behaviour and quality and found them agreeable enough to gather redundant data in long-term studies where participants occasionally fail to provide information from one modality [MCS+21].

Re-think the value of screen time as a metric Screen time as the sole metric to measure phone interaction could be unreliable, as applications might run in a foreground process, but users are not actively interacting with the phone. An example might be using the non-premium YouTube application to play music in the background while doing other activities.

Despite its easy obtainability, screen time as an indicator has already been criticised: "research community and society as a whole should move beyond the screen time debate and instead to examine the context and content surrounding social media use" [CRZ+20]. This might even be true for smartphone usage. An alternative to plain screen time has been used in this work: the ratio between screen time accumulated by using "problematic" applications to total screen time (see Section 4.1.6).

Using context-awareness capabilities provided by smartphones and smartwatches could help to find and generate a qualitative metric that captures the nature of phone interaction and its impact on the operator. Adding content analysis could prove to be a challenging topic already from a privacy point of view.

CHAPTER 6

Conclusion

6.1 Summary

This work presents a novel approach to supporting digital detox. It implements a simple intervention design using negative reinforcements in a smartphone-smartwatch ecosystem. At the time of writing this thesis, the use of wearables to support individuals in undergoing digital detox activities has not yet found wide applications, neither in the available literature nor in software solutions. A dedicated phone application collects phone usage data and sends daily statistics to a companion watch, displaying near-real-time usage information on its watch face. Additionally, users were able to define daily screen time limits for applications. Reaching the limit triggers a vibration on the watch that will remain until the user ceases to use this specific application. Furthermore, the phone application has access to physical health data from the watch (heart rate, steps and sleep cycles) via a Google Fit integration. Thus, the application can provide mash-up visualisation with data on daily phone usage and health-related data.

A small user study with six participants tested the prototype for two weeks. During the first week, participants significantly reduced their screen time, whereas, during the second week, screen time was reduced but not significantly. During the second week, the alarm feature was recommended to be employed. However, no significant reduction due to the alarm feature could be found. Further investigations revealed that participants either chose limits with lots of slack or altered limits to the maximum setting to use their applications a bit more without re-enabling their limit afterwards. This could either support the notion that participants could not define realistic limits for their desired behaviour or that the implemented design was too forceful, and participants chose to find ways to evade it.

Time analysis of phone screen time revealed that the usage of applications with a time limit increased between 7 pm and 7 am compared to the rest of the day and slightly reduced with each intervention week. With the available sleep data accessed, it has become apparent that the study's participants have been practising late bedtimes (around or after midnight) with accompanying phone usage beforehand. However, health data recordings were too sparse to deduct further insights other than the fact that participants' wake/sleep cycle did not meet the limit's time window assumptions.

The post-intervention interviews revealed several shortcomings of the proposed prototype and intervention designs. Despite all the issues, participants provided positive feedback for the watch face, thanks to its simplicity and ease of use. On the other hand, they reported a love-hate relationship towards the alarm feature. There was a significant reduction in phone screen time during the first week of the study in comparison to baseline usage, presumably due to wearing a watch as a physical reminder. This effect diminished during the second week as old behaviour patterns could again be observed, and even the introduction of the proposed alarm feature did not help reduce screen time further in some cases.

Thanks to the small size of the user study population ($n = 6$), the result's level of significance is debatable but instead should serve as an evaluation of a proof-of-concept design.

6.2 Smartwatches potential

Given the available results and the participants' experiences in this study, the proposed prototype setup has tested smartwatches' potential for digital detox and well-being. Participants reported increased awareness of their phone usage thanks to the watch face, the app, and some of their health metrics. However, given the short study duration, small population size and participant motivation, it's hard to derive anything substantial other than a positive first impression of the proposed prototype intervention.

Smartwatches could reduce direct interactions with a phone by providing easy access to required information [PBML16]. Simple information such as the time of day, the date, calendar entries, daily phone usage or controlling music is already available. Still, even more complex tasks are thinkable and might be available in the future. The capability of relaying notifications from phone to watch and whitelisting notifications applications to do so might be helpful when trying to reduce phone screen time — effectively taking a message delay approach [LLS+19]. Yet the impact of notifications on the attention cost and effective disruption [MPV+16, PVP18, SMY15] still remains. However, since this prototype did not monitor or collect notification events, there is no data to support this claim other than participants' reported experiences.

Despite being equally liked and hated by participants and having some encouraging first results, further adaptations to design and code-base are needed to tackle the identified issues and build better choice architecture.

6.2.1 Nudge Design

Rofarello and De Russis have already suggested using positive reinforcements [APS97], trigger events [SCB15] and the contextual awareness functionality of smartphones. With smartwatches' even further increased capabilities to identify context—from both the environment and internal sources—and more possibilities to interpret it, more engaging and precise nudge mechanics could be possible. With access to recorded sleep and activity data, making relevant suggestions or trigger events could be better timed.

6.3 Limitations

The presented findings have several limitations: First, due to the small study population, the implementation's prototypical state and the observed technical issues might have skewed the already scarce data. The relatively short duration of only two weeks additionally showed a diminishing effect of the intervention on overall screen time. The implemented intervention design allows us to see primarily short-term results, but with the alarm being particularly forceful, it can be expected to observe abandonment of the prototype over longer periods. Furthermore, the implemented intervention design does not strictly follow good practices from the present point of view in the literature; instead, it was chosen as being well adaptable to a watch interface and could be greatly improved to provide more tailored feedback to participants.

Considering that participants were all university students of similar age and socio-ecological descent and that participation earned them extra credit for coursework, the available data certainly holds some bias.

In some cases, usage events for participants (app usage and phone unlocks) had identical or very similar entries by app and duration properties but with different start and end timestamps. Since the phone's system time is usually set through the network provider, there seem to have been timestamp mismatches when usage stats got scraped. Suspicion is high that the observed phenomena occurred when participants possibly switched time zones or networks due to travelling, but the actual reasons remain hidden. This de-duplication process was pretty easy for usage events, thanks to having plenty of properties to distinguish between. Unfortunately, this was not the case for the unlock events. Henceforth, unlock event data has been excluded from further analysis and could not contribute to the findings.

From a usability point of view, the user interface relies heavily on dynamically chosen colours to help identify applications. Consequently, people with colour impairment will have difficulty reading the graphs or representation on the watch face.

6.3.1 Drawbacks

As with every technology, several factors need careful consideration and planning. Otherwise, good intentions can always backfire and create unintended side effects [Sch19].

Privacy and data protection

Usage and health data offer intimate insight into their owner's life, habits and personal preferences. Connecting multiple data sources further increases the level of detail and could help to foster healthier lifestyles. However, in the same manner, this level makes it attractive for companies, corporations, governments and criminals for various reasons: marketing, surveillance and exploits or monetisation of personal shortcomings. Recent laws like the EU's (DSGVO)[1] introduced law-bound guidelines and high fines for violations of those principles in hopes of strengthening data protection.

Critics of these data-gathering practices are probably right. Yet, given the possibilities offered by insights from this data to create a chance for a healthier life, it might be worth taking a calculated risk.

Increased connected-ness and digital entanglement

Adding another piece of technology designed to support smartphone interaction can be seen as truthfully troublesome. Without proper care, the watch could help make the connection between the phone and the user more vital and difficult to evade or break. Although Pizza et al. suspect that a smartwatch, and depending on the activity that's undertaken, the effect of disruptions through notification might be neglectable [PBML16].

6.4 Future outlook

Besides having identified several shortcomings of the proposed prototype and possible ways to fix them, there is another problem with the available implementation: Google Fit is deprecated and will only receive support until the end of 2024. Google is pushing a new API called *Health Connect*[2], which supposedly offers more privacy control over on-device health and fitness data. Implementing Watchful using Health Connect could grow the number of compatible smartwatches and the number of users that could be supported on their digital detox journey.

Furthermore, to help deal with data gaps, implementing alternative or complementing data-recording modalities could prove beneficial. For instance, "tappigraphy" [BHG19, BG18] could enhance screen-time metrics with interaction data, whereas smartphone sensors, in turn, could complement wearable recordings, as has been demonstrated by Massar et al. [MCS+21]. In the case of *Android*, APIs are available to gather step data through phone-integrated motion sensors and even get callbacks on significant motion events[3].

[1] DSGVO, `https://www.datenschutz-grundverordnung.eu/`–Accessed August 2023

[2] Google Health Connect, `https://developer.android.com/guide/health-and-fitness/health-connect` — Accessed August 2023

[3] Android Motion sensors, `https://developer.android.com/develop/sensors-and-location/sensors/sensors_motion` — Accessed January 2024

6.4.1 Improved *Watchful* design

No prototype is perfect, and based on the participant feedback (see Section 4.3.9) and insights from the studies results and observed behaviour (see Section 4.1.5 and D), the following improved prototype design could be a possibility for further research. Besides, general guidelines recommended in the literature(presented in Section 2.8) should be further revised and taken into account to design future intervention mechanics.

The following bullet points collect observations and participant's suggestions:

Alarm intervention: From a technical perspective, the prototype's implementation worked fine within the scope of this thesis. However, several improvements have been suggested to reduce friction in the user experience and to make the mechanism more reliable:

- choosing alarm limits could respect past week usage and recommend a realistic value or even automatically adapt the value over time.
- Progress on reaching limits should be easily visible.
- The alarm feature should be less forceful with the possibility of snoozing it.
- While an alarm is active, closing the application responsible for triggering the alarm could turn off the alarm automatically.
- In case of an unavailable watch, a fallback alarm on the phone should take over.

Watch face: Minor convenience features could make the watch face interventions more useful for a wider audience without using the provided dedicated *Watchful* watch face.

- WearOS supports tiles and complication implementations besides watch faces to make information more easily available in a customisable way for personal preferences.
- Stale usage data needs to be detected, and the user informed about missing fresh data.

Sleep data: With access to more reliable sleep activity data could offer additional interventions:

- With access to bed- and wake-time, phone usage before bed could be detected and discouraged.
- Late-night phone usage could be discouraged through an additional intervention, and alarm limits time windows could better align with wake/sleep cycles.
- Include the impact of smartphone usage on sleep quality as already proposed by Arora et al. [ACB21].

Further technical improvements: The proposed prototype suffers from a very limited range of supported watch models due to its reliance on *Google Fit* to capture health and sleep data. Unfortunately, this Google service has become deprecated as of 2024 and should be migrated to *Health Connect*[4]. Removal of the *Google Fit* dependency would make the prototype further applicable to a greater number of watch models and thus allow it to be used by a wider user base.

In addition, redundant modalities to record phone usage, interaction metrics and physiological/activity data could greatly improve the reliability of the proposed interventions. Finding new metrics to describe phone interaction beyond a quantitative perspective could further introduce new meaning to "Digital Detox" and make for a better experience for the audience.

Guidance, appraisal and reward Providing users with an appraisal of their progress and highlighting observed changes could encourage long-term habit formation. However, finding the right words and rewards that motivate users are additional challenges.

In contrast, providing better guidance on phone usage through companion devices could be easily done, resulting hopefully in a better user experience and supporting habit formation.

6.5 Closing words

This thesis has investigated the application of a smartwatch for digital detox intervention to decrease phone screen time, which had no published or similar implementation at the time. The results showed promising first impressions and good feedback from participants, but they also revealed a number of shortcomings and limitations. The small study population and short duration prohibit deriving generalised recommendations for long-term effects.

Working on the thesis made the complex nature of the relationship we have developed with our smartphones even more apparent to me. The sheer amount of continuous work and motivation necessary to adopt personal habits for healthier versions is incredible and has been greatly underestimated by me.

Two years have passed from the original idea to the conclusion of the write-up, during which time the prototype has grown into its latest form. Yet, despite (or perhaps because of) the long time spent in development, a myriad of issues arose with the prototype when flung into the real world. In retrospect, I would have done many things differently, in some cases faster and in others more deliberately. However, isn't that almost always the realisation afterwards?

Nothing is easy; The journey is the reward; Bad roads and uncharted regions are part of it;

[4] *Google Fit* migration guide, `https://developer.android.com/health-and-fitness/guides/health-connect/migrate/migration-guide` – Accessed February 2024

Let's cherish the experience and hope these notes and directions might help someone in the future besides myself.

List of Figures

1.1 *Watchful* logo 1
1.2 *Watchful*'s watch face in action 5
1.3 Research Design (qualitative data in *blue*, quantitative data in *green* and mixed data in *pink*) 7

2.1 Google search trends for the term 'digital detox' 10
2.2 Feature and features clusters commonly used in detox applications [MRDR19, LLS+19] 12
2.3 Acceptance/rejections factors according to Almourad et al. [AAS+21] . . 15
2.4 *AppDetox* [LBG13] 16
2.5 *MyTime* by Hiniker et al. [HHKK16] 17
2.6 iOS's *Screen Time* and *Focus* (iOS 16) 17
2.7 Android's *Digital Wellbeing* (Android 13) 18
2.8 screenshots of *GoodVibe* by Okeke et al. [OSDE18] and *GoalKeeper* by Kim et al. [KJKL19] 19
2.9 Screenshots of *GoalKeeper* [KJKL19] 20
2.10 *Socialize* by Monge Rofarello & De Russis [MRDR19] 21
2.11 Screenshots of *Nudgit* [SMWM22] 22
2.12 An intervention design by Purohit et al. using iOS' *Shortcuts* app [PBS+23] 23
2.13 Smart lamps assisting building better sleep hygiene [LLKC17, CLPK19] . 24
2.14 *TIME OFF!* a design study by *RSW* 25
2.15 Learning procedures for activity recognition. [WCY19] 26
2.16 Architecture of a wearable medical sensors system [MSKRJ17] 27

3.1 Watch face example and early sketch 34
3.2 Early draft of the "alarm nudge"-intervention. 35
3.3 Colour consistency per application across views and devices. 37
3.4 User interface of *Watchful* 39
3.5 User interface of *Watchful* 41
3.6 Health data visualisations for the last 24 hours. 41
3.7 Explanation of the available fields of the *Watchful* watch face that was supplied to the participants. 42
3.8 *Watchful*'s watch application screens 43
3.9 Data flow between all components. 44

3.10 Granting required permissions for *Watchful* 45
3.11 *Watchful*'s watch alarm in action 46
3.12 User study design 47
3.13 Transforming descriptive event data to data points for easier analysis. . . 49
3.14 *Grafana* dashboards for data exploration. 50

4.1 Use of *Watchful* 53
4.2 Usage data group analysis 54
4.3 Weekly overall usage box-plots 55
4.4 box-plots for weekly overall usage for applications with specified limits (kept until the end of the study). 55
4.5 box-plots for the weekly count of used applications. 56
4.6 average screen time usage per hour per study phase (95% confidence interval of null-value cleaned in colour, complete dataset in grey) 57
4.7 average screen time usage per hour per study phase in daily and nightly splits (95% confidence interval of null-value cleaned in colour, complete dataset in grey) 57
4.8 average screen time usage per hour per study phase (95% confidence interval of null-value cleaned in colour, complete dataset in grey) 59
4.9 average screen time usage per hour per study phase in daily and nightly splits (95% confidence interval of null-value cleaned in colour, complete dataset in grey) 59
4.10 Exemplary plot of combined usage and sleep data 61
4.11 Participant 1' sleep 62
4.12 Participant 2' sleep 62
4.13 Participant 3' sleep 63
4.14 Participant 4' sleep 63
4.15 Participant 5' sleep 64
4.16 Participant 6' sleep 64

5.1 Participant 2's usage of *com.instagram.android* and *net.lovoo.android* . . 78
5.2 Participant 1's usage of *com.instagram.android* and *com.android.chrome* . 78

A.1 Survey results from Likert scales questions from all participants ($n = 9$) . 111

C.1 Participant 1's top 5 most used applications for each week of the study . 116
C.2 Participant 2's top 5 most used applications for each week of the study . 117
C.3 Participant 3's top 5 most used applications for each week of the study . 118
C.4 Participant 4's top 5 most used applications for each week of the study . 119
C.5 Participant 5's top 5 most used applications for each week of the study . 120
C.6 Participant 6s's top 5 most used applications for each week of the study . 121

D.1 Participant 1's usage of *com.instagram.android* and *com.android.chrome* 123
D.2 Participant 2's usage of *com.instagram.android* and *net.lovoo.android* . . 124

D.3 Participant 3's usage of *com.google.android.youtube* and *com.instagram.android* 124
D.4 Participant 4's usage of *com.zhiliaoapp.musically/TikTok* 124
D.5 Participant 5's usage of *com.android.chrome/Google Chrome* 124
D.6 Participant 6's usage of *com.amazon.avod.thirdpartyclient*, *com.instagram.android* and *com.vanced.android.youtube* 125

List of Tables

A.1 Survey questions — Part 1 109
A.2 Survey questions — Part 2 110
A.3 Survey questions — part 3 110

C.1 Participant 1's weekly top 5 application totals 116
C.2 Participant 2's weekly top 5 application totals 117
C.3 Participant 3's weekly top 5 application totals 118
C.4 Participant 4's weekly top 5 application totals 119
C.5 Participant 5's weekly top 5 application totals 120
C.6 Participant 6's weekly top 5 application totals 121

E.1 Hourly screentime including null values 128
E.2 Hourly screentime including null values 129
E.3 Hourly screentime cleaned from null values 130
E.4 Hourly screentime cleaned from null values 131
E.5 Hourly screentime of "limited" applications including null values 132
E.6 Hourly screentime of "limited" applications including null values 133
E.7 Hourly proportion of "limited" apps'screen time of total screentime including null values 134
E.8 Hourly proportion of "limited" apps'screen time of total screentime including null values 135
E.9 Hourly proportion of "limited" apps'screen time of total screentime cleaned from null values 136
E.10 Hourly proportion of "limited" apps'screen time of total screentime cleaned from null values 137

Bibliography

[AAS+21] Mohamed Basel Almourad, Amen Alrobai, Tiffany Skinner, Mohammed Hussain, and Raian Ali. Digital wellbeing tools through users lens. *Technology in Society*, 67:101778, 2021.

[ACB21] Anshika Arora, Pinaki Chakraborty, and MPS Bhatia. Intervention of wearables and smartphones in real time monitoring of sleep and behavioral health: an assessment using adaptive neuro-fuzzy technique. *Arabian Journal for Science and Engineering*, pages 1–26, 2021.

[ACB22] Anshika Arora, Pinaki Chakraborty, and MPS Bhatia. Intelligent behavioral health analysis in youth using real-time smartphone and wearable data. In *Proceedings of International Conference on Communication and Computational Technologies: ICCCT 2022*, pages 653–666. Springer, 2022.

[Ame13] Morgan G Ames. Managing mobile multitasking: The culture of iphones on stanford campus. In *Proceedings of the 2013 conference on Computer supported cooperative work*, pages 1487–1498, 2013.

[APS97] Henk Aarts, Theo Paulussen, and Herman Schaalma. Physical exercise habit: on the conceptualization and formation of habitual health behaviours. *Health education research*, 12(3):363–374, 1997.

[BG18] M Balerna and A Ghosh. The details of past actions on a smartphone touchscreen are reflected by intrinsic sensorimotor dynamics. digital med, 7, article 4, 2018.

[BHG19] Jay N Borger, Reto Huber, and Arko Ghosh. Capturing sleep–wake cycles by using day-to-day smartphone touchscreen interactions. *NPJ digital medicine*, 2(1):73, 2019.

[BM21] Peter André Busch and Stephen McCarthy. Antecedents and consequences of problematic smartphone use: A systematic literature review of an emerging research area. *Computers in human behavior*, 114:106414, 2021.

[BMLF+15] Joël Billieux, Pierre Maurage, Olatz Lopez-Fernandez, Daria J Kuss, and Mark D Griffiths. Can disordered mobile phone use be considered a behavioral addiction? an update on current evidence and a comprehensive model for future research. *Current Addiction Reports*, 2(2):156–162, 2015.

[BWO+19] Katie-Jane Brickwood, Greig Watson, Jane O'Brien, Andrew D Williams, et al. Consumer-based wearable activity trackers increase physical activity participation: systematic review and meta-analysis. *JMIR mHealth and uHealth*, 7(4):e11819, 2019.

[CBK10] Kaustav Chakraborty, Debashis Basu, and KG Vijaya Kumar. Internet addiction: consensus, controversies, and the way ahead, 2010.

[CCB15] Marta E Cecchinato, Anna L Cox, and Jon Bird. Smartwatches: the good, the bad and the ugly? In *Proceedings of the 33rd Annual ACM Conference extended abstracts on human factors in computing systems*, pages 2133–2138, 2015.

[CEF+12] Marie Chan, Daniel Estève, Jean-Yves Fourniols, Christophe Escriba, and Eric Campo. Smart wearable systems: Current status and future challenges. *Artificial intelligence in medicine*, 56(3):137–156, 2012.

[CLPK19] Ha-yeona Choi, Kyung-Ryonga Lee, Young-Wooa Park, and Cha-joong Kim. Designing for improving sleep hygiene through the reflection of smartphone awareness with ambient lighting. In *International Association of Societies of Design Research Conference*. Manchester Metropolitan University, 2019.

[CML09] Sunny Consolvo, David W McDonald, and James A Landay. Theory-driven design strategies for technologies that support behavior change in everyday life. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 405–414, 2009.

[CRZ+20] Sarah M Coyne, Adam A Rogers, Jessica D Zurcher, Laura Stockdale, and McCall Booth. Does time spent using social media impact mental health?: An eight year longitudinal study. *Computers in human behavior*, 104:106160, 2020.

[DRS18] Jessilyn Dunn, Ryan Runge, and Michael Snyder. Wearables and the medical revolution. *Personalized medicine*, 15(5):429–448, 2018.

[FPE10] Richard R Fletcher, Ming-Zher Poh, and Hoda Eydgahi. Wearable sensors: Opportunities and challenges for low-cost health care. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pages 1763–1766. IEEE, 2010.

[Gar15] Benjamin Gardner. 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.

[GLAO23] Natalie Gold, Yiling Lin, Richard Ashcroft, and Magda Osman. 'better off, as judged by themselves': do people support nudges as a method to change their own behavior? *Behavioural Public Policy*, 7(1):25–54, 2023.

[GMvO+18] Roxanne Gal, Anne M May, Elon J van Overmeeren, Monique Simons, and Evelyn M Monninkhof. The effect of physical activity interventions comprising wearables and smartphone applications on physical activity: a systematic review and meta-analysis. *Sports medicine-open*, 4:1–15, 2018.

[Gof16] Erving Goffman. The presentation of self in everyday life. In *Social Theory Re-Wired*, pages 482–493. Routledge, 2016.

[GT20] Kerry Gibson and Susanna Trnka. Young people's priorities for support on social media:"it takes trust to talk about these issues". *Computers in human behavior*, 102:238–247, 2020.

[HHKK16] Alexis Hiniker, Sungsoo Hong, Tadayoshi Kohno, and Julie A Kientz. Mytime: designing and evaluating an intervention for smartphone non-use. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, pages 4746–4757, 2016.

[HM19] Dennis Hummel and Alexander Maedche. How effective is nudging? a quantitative review on the effect sizes and limits of empirical nudging studies. *Journal of Behavioral and Experimental Economics*, 80:47–58, 2019.

[IMD+21] Sheikh MA Iqbal, Imadeldin Mahgoub, E Du, Mary Ann Leavitt, and Waseem Asghar. Advances in healthcare wearable devices. *NPJ Flexible Electronics*, 5(1):9, 2021.

[ISVO20] Netta Iivari, Sumita Sharma, and Leena Ventä-Olkkonen. Digital transformation of everyday life–how covid-19 pandemic transformed the basic education of the young generation and why information management research should care? *International journal of information management*, 55:102183, 2020.

[JFHZ05] Nassim Jafarinaimi, Jodi Forlizzi, Amy Hurst, and John Zimmerman. Breakaway: an ambient display designed to change human behavior. In *CHI'05 extended abstracts on Human factors in computing systems*, pages 1945–1948, 2005.

[JG03] Eric J Johnson and Daniel Goldstein. Do defaults save lives?, 2003.

[JKC16] Yoonhyuk Jung, Seongcheol Kim, and Boreum Choi. Consumer valuation of the wearables: The case of smartwatches. *Computers in Human Behavior*, 63:899–905, 2016.

[JSD+12] Eric J Johnson, Suzanne B Shu, Benedict GC Dellaert, Craig Fox, Daniel G Goldstein, Gerald Häubl, Richard P Larrick, John W Payne, Ellen Peters, David Schkade, et al. Beyond nudges: Tools of a choice architecture. *Marketing letters*, 23:487–504, 2012.

[Kah11] Daniel Kahneman. *Thinking, fast and slow.* Macmillan, 2011.

[KG11] Daria J. Kuss and Mark D. Griffiths. Online social networking and addiction—a review of the psychological literature. *International Journal of Environmental Research and Public Health*, 8(9):3528–3552, 2011.

[KI13] Holly Korda and Zena Itani. Harnessing social media for health promotion and behavior change. *Health promotion practice*, 14(1):15–23, 2013.

[KJKL19] Jaejeung Kim, Hayoung Jung, Minsam Ko, and Uichin Lee. Goalkeeper: exploring interaction lockout mechanisms for regulating smartphone use. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 3(1):1–29, 2019.

[KPK+19] Seung-Gon Kim, Jong Park, Hun-Tae Kim, Zihang Pan, Yena Lee, and Roger S McIntyre. The relationship between smartphone addiction and symptoms of depression, anxiety, and attention-deficit/hyperactivity in south korean adolescents. *Annals of general psychiatry*, 18(1):1–8, 2019.

[KWB18] Geza Kovacs, Zhengxuan Wu, and Michael S Bernstein. Rotating online behavior change interventions increases effectiveness but also increases attrition. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):1–25, 2018.

[Lan84] David S Landes. Revolution in time: clocks and the making of the modern world. *Revolution in time: clocks and the making of the modern world.. DS Landes. Harvard University Press*, 1984.

[LBG13] Markus Löchtefeld, Matthias Böhmer, and Lyubomir Ganev. Appdetox: helping users with mobile app addiction. In *Proceedings of the 12th international conference on mobile and ubiquitous multimedia*, pages 1–2, 2013.

[LLKC17] JongSung Lee, Ji Yeon Lee, Seog Wang Kim, and Jun Dong Cho. D-tox: Inducing digital detox for nighttime via smart lamp applied gamification. In *Proceedings of the 2017 Conference on Interaction Design and Children*, pages 497–502, 2017.

[LLS+19] Ulrik Lyngs, Kai Lukoff, Petr Slovak, Reuben Binns, Adam Slack, Michael Inzlicht, Max Van Kleek, and Nigel Shadbolt. Self-control in cyberspace: Applying dual systems theory to a review of digital self-control tools. In *proceedings of the 2019 CHI conference on human factors in computing systems*, pages 1–18, 2019.

[LVJPW10] Phillippa Lally, Cornelia HM Van Jaarsveld, Henry WW Potts, and Jane Wardle. How are habits formed: Modelling habit formation in the real world. *European journal of social psychology*, 40(6):998–1009, 2010.

[Mar02] Thomas L Martin. Time and time again: Parallels in the development of the watch and the wearable computer, 2002.

[MBL+17] Donald McMillan, Barry Brown, Airi Lampinen, Moira McGregor, Eve Hoggan, and Stefania Pizza. Situating wearables: Smartwatch use in context. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, page 3582–3594, New York, NY, USA, 2017. Association for Computing Machinery.

[MC15] Vivian Genaro Motti and Kelly Caine. Users' privacy concerns about wearables: impact of form factor, sensors and type of data collected. In *Financial Cryptography and Data Security: FC 2015 International Workshops, BITCOIN, WAHC, and Wearable, San Juan, Puerto Rico, January 30, 2015, Revised Selected Papers*, pages 231–244. Springer, 2015.

[MCS+21] Stijn AA Massar, Xin Yu Chua, Chun Siong Soon, Alyssa SC Ng, Ju Lynn Ong, Nicholas IYN Chee, Tih Shih Lee, Arko Ghosh, and Michael WL Chee. Trait-like nocturnal sleep behavior identified by combining wearable, phone-use, and self-report data. *NPJ digital medicine*, 4(1):90, 2021.

[MHF22] Ludivine Martin, Laetitia Hauret, and Chantal Fuhrer. Digitally transformed home office impacts on job satisfaction, job stress and job productivity. covid-19 findings. *Plos one*, 17(3):e0265131, 2022.

[MHM16] Abhinav Mehrotra, Robert Hendley, and Mirco Musolesi. Towards multi-modal anticipatory monitoring of depressive states through the analysis of human-smartphone interaction. In *Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing: adjunct*, pages 1132–1138, 2016.

[MHZ+19] Xi Mei, Zhenyu Hu, Dongsheng Zhou, Qi Zhou, Xingxing Li, Xiaojia Wang, and Pan Jing. Sleep patterns, mobile phone use and psychological symptoms among adolescents in coastal developed city of china: an exploratory cross-sectional study. *Sleep and Biological Rhythms*, 17:233–241, 2019.

[Mil23] Stuart Mills. Nudge/sludge symmetry: on the relationship between nudge and sludge and the resulting ontological, normative and transparency implications. *Behavioural Public Policy*, 7(2):309–332, 2023.

[MMMMY14] Leanne Morrison, Rona Moss-Morris, Susan Michie, and Lucy Yardley. Optimizing engagement with i nternet-based health behaviour change interventions: Comparison of self-assessment with and without tailored feedback using a mixed methods approach. *British journal of health psychology*, 19(4):839–855, 2014.

[MPV+16] Abhinav Mehrotra, Veljko Pejovic, Jo Vermeulen, Robert Hendley, and Mirco Musolesi. My phone and me: understanding people's receptivity to mobile notifications. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, pages 1021–1032, 2016.

[MRDR19] Alberto Monge Roffarello and Luigi De Russis. The race towards digital wellbeing: Issues and opportunities. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pages 1–14, 2019.

[MRDR23] Alberto Monge Roffarello and Luigi De Russis. Achieving digital wellbeing through digital self-control tools: A systematic review and meta-analysis. *ACM Transactions on Computer-Human Interaction*, 30(4):1–66, 2023.

[MRJ+13] Susan Michie, Michelle Richardson, Marie Johnston, Charles Abraham, Jill Francis, Wendy Hardeman, Martin P Eccles, James Cane, and Caroline E Wood. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Annals of behavioral medicine*, 46(1):81–95, 2013.

[MSKRJ17] Arsalan Mosenia, Susmita Sur-Kolay, Anand Raghunathan, and Niraj K Jha. Wearable medical sensor-based system design: A survey. *IEEE Transactions on Multi-Scale Computing Systems*, 3(2):124–138, 2017.

[MVS16] Robert Münscher, Max Vetter, and Thomas Scheuerle. A review and taxonomy of choice architecture techniques. *Journal of Behavioral Decision Making*, 29(5):511–524, 2016.

[NMA12] Michael I Norton, Daniel Mochon, and Dan Ariely. The ikea effect: When labor leads to love. *Journal of consumer psychology*, 22(3):453–460, 2012.

[NWQ06] David T Neal, Wendy Wood, and Jeffrey M Quinn. Habits—a repeat performance. *Current directions in psychological science*, 15(4):198–202, 2006.

[ORMR12] Antti Oulasvirta, Tye Rattenbury, Lingyi Ma, and Eeva Raita. Habits make smartphone use more pervasive. *Personal and Ubiquitous computing*, 16:105–114, 2012.

[OSDE18] Fabian Okeke, Michael Sobolev, Nicola Dell, and Deborah Estrin. Good vibrations: can a digital nudge reduce digital overload? In *Proceedings of the 20th international conference on human-computer interaction with mobile devices and services*, pages 1–12, 2018.

[PB09] Alexandros Pantelopoulos and Nikolaos G Bourbakis. A survey on wearable sensor-based systems for health monitoring and prognosis. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(1):1–12, 2009.

[PBH20] Aditya Kumar Purohit, Louis Barclay, and Adrian Holzer. Designing for digital detox: Making social media less addictive with digital nudges. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–9, 2020.

[PBML16] Stefania Pizza, Barry Brown, Donald McMillan, and Airi Lampinen. Smartwatch in vivo. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 5456–5469, 2016.

[PBS+23] Aditya Kumar Purohit, Torben Jan Barev, Sofia Schöbel, Andreas Janson, and Adrian Holzer. Designing for digitalwellbeing on a smartphone: Co-creation of digital nudges to mitigate instagram overuse. In T.X. Bui, editor, *HICSS 2023: Proceedings of the 56th Hawaii International Conference on System Sciences*, pages 4087–4096. Honolulu: ScholarSpace, 2023.

[PH19] Aditya Kumar Purohit and Adrian Holzer. Functional digital nudges: Identifying optimal timing for effective behavior change. In *Extended abstracts of the 2019 CHI conference on human factors in computing systems*, pages 1–6, 2019.

[PH21] Aditya Kumar Purohit and Adrian Holzer. Unhooked by design: Scrolling mindfully on social media by automating digital nudges. In *AMCIS*, volume 21, pages 1–10, 2021.

[PlM+23] Douglas A. Parry, Daniel B. le Roux, Joshua Morton, Ruth Pons, Ruan Pretorius, and Andre Schoeman. Digital wellbeing applications: Adoption, use and perceived effects. *Computers in Human Behavior*, 139:107542, 2023.

[PSS20] Niels Peek, Mark Sujan, and Philip Scott. Digital health and care in pandemic times: impact of covid-19. *BMJ Health & Care Informatics*, 27(1), 2020.

[PVCB18] Charlie Pinder, Jo Vermeulen, Benjamin R Cowan, and Russell Beale. Digital behaviour change interventions to break and form habits. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 25(3):1–66, 2018.

[PVP18] Martin Pielot, Amalia Vradi, and Souneil Park. Dismissed! a detailed exploration of how mobile phone users handle push notifications. In *Proceedings of the 20th international conference on human-computer interaction with mobile devices and services*, pages 1–11, 2018.

[PZ17] Alfredo J Perez and Sherali Zeadally. Privacy issues and solutions for consumer wearables. *It Professional*, 20(4):46–56, 2017.

[RAS+21] Theda Radtke, Theresa Apel, Konstantin Schenkel, Jan Keller, and Eike von Lindern. Digital detox: An effective solution in the smartphone era? a systematic literature review. *Mobile Media & Communication*, page 20501579211028647, 2021.

[RPP14] Reza Rawassizadeh, Blaine A Price, and Marian Petre. Wearables: Has the age of smartwatches finally arrived? *Communications of the ACM*, 58(1):45–47, 2014.

[SAACP18] LH Segura Anaya, Abeer Alsadoon, Nectar Costadopoulos, and PWC Prasad. Ethical implications of user perceptions of wearable devices. *Science and engineering ethics*, 24:1–28, 2018.

[SAVL+16] Stephanie Schoeppe, Stephanie Alley, Wendy Van Lippevelde, Nicola A Bray, Susan L Williams, Mitch J Duncan, and Corneel Vandelanotte. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, 13(1):1–26, 2016.

[SCB14] Katarzyna Stawarz, Anna L Cox, and Ann Blandford. Don't forget your pill! designing effective medication reminder apps that support users' daily routines. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2269–2278, 2014.

[SCB15] Katarzyna Stawarz, Anna L Cox, and Ann Blandford. Beyond self-tracking and reminders: designing smartphone apps that support habit formation. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, pages 2653–2662, 2015.

[Sch19] RX Schwartz. Ulysses' ropes and the inherent limits of digital self-control tools. *5th International Conference on the History and Philosophy of Computing*, 2019.

[SCT+19] Thomas Svensson, Ung-il Chung, Shinichi Tokuno, Mitsuteru Nakamura, and Akiko Kishi Svensson. A validation study of a consumer wearable sleep tracker compared to a portable eeg system in naturalistic conditions. *Journal of Psychosomatic Research*, 126:109822, 2019.

[SDVJ+19] Jussi Seppälä, Ilaria De Vita, Timo Jämsä, Jouko Miettunen, Matti Isohanni, Katya Rubinstein, Yoram Feldman, Eva Grasa, Iluminada Corripio, Jesus Berdun, et al. Mobile phone and wearable sensor-based mhealth approaches for psychiatric disorders and symptoms: systematic review. *JMIR mental health*, 6(2):e9819, 2019.

[SH16] Maya Samaha and Nazir S Hawi. Relationships among smartphone addiction, stress, academic performance, and satisfaction with life. *Computers in human behavior*, 57:321–325, 2016.

[SHG22] Tucker Stuart, Jessica Hanna, and Philipp Gutruf. Wearable devices for continuous monitoring of biosignals: Challenges and opportunities. *APL bioengineering*, 6(2), 2022.

[SHN+17] Suranga Seneviratne, Yining Hu, Tham Nguyen, Guohao Lan, Sara Khalifa, Kanchana Thilakarathna, Mahbub Hassan, and Aruna Seneviratne. A survey of wearable devices and challenges. *IEEE Communications Surveys & Tutorials*, 19(4):2573–2620, 2017.

[SMRDRA21] RX Schwartz, Alberto Monge Roffarello, Luigi De Russis, and Panagiotis Apostolellis. 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*, pages 1–7, 2021.

[SMWM22] Ambika Shahu, Alexander Melem, Philipp Wintersberger, and Florian Michahelles. Nudgit-reducing online news consumption by digital nudges. In *Adjunct Publication of the 24th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pages 1–5, 2022.

[SMY15] Cary Stothart, Ainsley Mitchum, and Courtney Yehnert. The attentional cost of receiving a cell phone notification. *Journal of experimental psychology: human perception and performance*, 41(4):893, 2015.

[SP13] Akane Sano and Rosalind W Picard. Stress recognition using wearable sensors and mobile phones. In *2013 Humaine association conference on affective computing and intelligent interaction*, pages 671–676. IEEE, 2013.

[SP19] Jini Jojo Stephen and P Prabu. Detecting the magnitude of depression in twitter users using sentiment analysis. *International Journal of Electrical and Computer Engineering*, 9(4):3247, 2019.

[SPA+15] Akane Sano, Andrew J Phillips, Z Yu Amy, Andrew W McHill, Sara Taylor, Natasha Jaques, Charles A Czeisler, Elizabeth B Klerman, and Rosalind W Picard. Recognizing academic performance, sleep quality, stress level, and mental health using personality traits, wearable sensors and mobile phones. In *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, pages 1–6. IEEE, 2015.

[SSL10] Daniel Scanfeld, Vanessa Scanfeld, and Elaine L Larson. Dissemination of health information through social networks: Twitter and antibiotics. *American journal of infection control*, 38(3):182–188, 2010.

[ST03] Cass R Sunstein and Richard H Thaler. Libertarian paternalism is not an oxymoron. *The University of Chicago Law Review*, pages 1159–1202, 2003.

[Tha18] Richard H Thaler. Nudge, not sludge, 2018.

[TMM+20] Matilda Swee Sun Tang, Katherine Moore, Andrew McGavigan, Robyn A Clark, and Anand N Ganesan. Effectiveness of wearable trackers on physical activity in healthy adults: systematic review and meta-analysis of randomized controlled trials. *JMIR mHealth and uHealth*, 8(7):e15576, 2020.

[TS08] Richard H Thaler and Cass R Sunstein. *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press, 2008.

[Udd19] Md Zia Uddin. A wearable sensor-based activity prediction system to facilitate edge computing in smart healthcare system. *Journal of Parallel and Distributed Computing*, 123:46–53, 2019.

[VSVB+17] Aku Visuri, Zhanna Sarsenbayeva, Niels Van Berkel, Jorge Goncalves, Reza Rawassizadeh, Vassilis Kostakos, and Denzil Ferreira. Quantifying sources and types of smartwatch usage sessions. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 3569–3581, 2017.

[VZC20] Gabriela Villalobos-Zúñiga and Mauro Cherubini. Apps that motivate: A taxonomy of app features based on self-determination theory. *International Journal of Human-Computer Studies*, 140:102449, 2020.

[WCY19] Yan Wang, Shuang Cang, and Hongnian Yu. A survey on wearable sensor modality centred human activity recognition in health care. *Expert Systems with Applications*, 137:167–190, 2019.

[Wei76] Joseph Weizenbaum. *Computer power and human reason: From judgment to calculation*. WH Freeman & Co, 1976.

[WI23] Michael Wadsley and Niklas Ihssen. Restricting social networking site use for one week produces varied effects on mood but does not increase explicit or implicit desires to use snss: Findings from an ecological momentary assessment study. *PLoS one*, 18(11):e0293467, 2023.

[Wid20] Kelly Widdicks. When the good turns ugly: Speculating next steps for digital wellbeing tools. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*, pages 1–6, 2020.

[WR16] Wendy Wood and Dennis Rünger. Psychology of habit. *Annual Review of Psychology*, 67(1):289–314, 2016. PMID: 26361052.

[WSB16] Markus Weinmann, Christoph Schneider, and Jan vom Brocke. Digital nudging. *Business & Information Systems Engineering*, 58:433–436, 2016.

[WSRS22] Patrick Wilms, Jan Schröder, Rüdiger Reer, and Lorenz Scheit. The impact of "home office" work on physical activity and sedentary behavior during the covid-19 pandemic: a systematic review. *International Journal of Environmental Research and Public Health*, 19(19):12344, 2022.

[YFLL20] Jiaxin Yang, Xi Fu, Xiaoli Liao, and Yamin Li. Association of problematic smartphone use with poor sleep quality, depression, and anxiety: A systematic review and meta-analysis. *Psychiatry Research*, 284:112686, 2020.

APPENDIX A

Survey questions

Q1	How much did your awareness about your phone usage increased during this first week? → 5 point Likert-scale (*not at all* to *very much*) see Figure A.1a
Q2	To what degree were you aware of the watch? → 5 point Likert-scale (*not at all* to *very much*) see Figure A.1b
Q3	How much did the watch bother you? → 5 point Likert-scale (*not at all* to *very much*) see Figure A.1c
Q4	Which app did you check for your daily screen time more often during this first week? see Figure A.1d
Q5	How do you feel about reading the visualisations available in the Watchful phone app? → 5 point Likert-scale (*not at all* to *very much*) see Figure A.1e

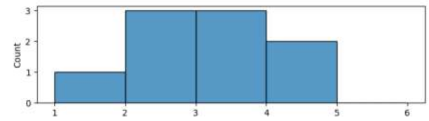
Table A.1: Survey questions — Part 1

Q6	How much value do the visualisations available in the Watchful phone app offer to you? → 5 point Likert-scale (*nothing* to *a lot*) see Figure A.1f
Q7	How much value does the available health data in combination with your phone usage data offer? → 5 point Likert-scale (*nothing* to *a lot*) see Figure A.1g
Q8	Do you have anything to add about your experiences with the phone app so far?

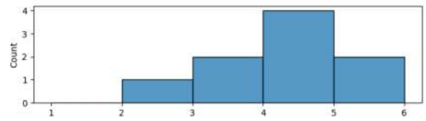
Table A.2: Survey questions — Part 2

Q9	How do you feel about interpreting the watch face? → 5 point Likert-scale (*difficult to read* to *easy to read*) see Figure A.1h
Q10	How often did you use the Watchful watch app directly? → 5 point Likert-scale (*not at all* to *very often*) see Figure A.1i
Q11	Is there something that you want to tell us about your experiences with the watch or watch-app?
Q12	How useful are the available statistics to you? → 5 point Likert-scale (*not at all* to *a lot*) see Figure A.1j
Q13	Are there any changes to the available statistics that you would recommend?

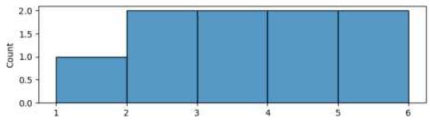
Table A.3: Survey questions — part 3



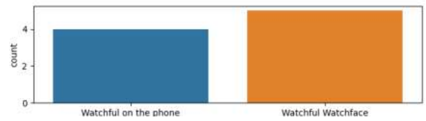
(a) Survey **Q1** responses



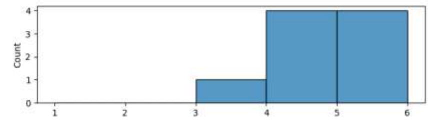
(b) Survey **Q2** responses



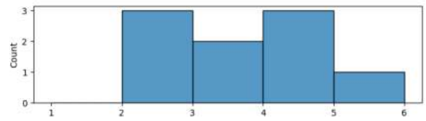
(c) Survey **Q3** responses



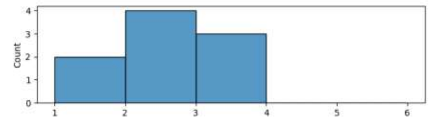
(d) Survey **Q4** responses



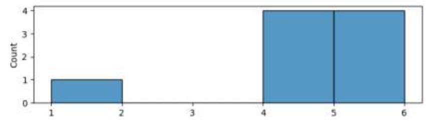
(e) Survey **Q5** responses



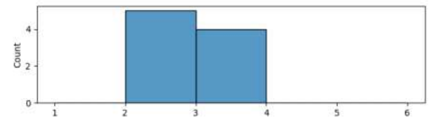
(f) Survey **Q6** responses



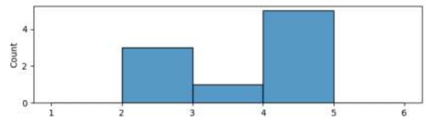
(g) Survey **Q7** responses



(h) Survey **Q9** responses



(i) Survey **Q10** responses



(j) Survey **Q12** responses

Figure A.1: Survey results from Likert scales questions from all participants ($n = 9$)

APPENDIX B

Interview guide

We appreciate you took the time to participate in our study.

Regarding your overall experience with our digital detoxification prototype, we have a few questions for you. As always, there is nothing you can do wrong! There are no right or wrong responses! We're interested in learning about your personal experience.

We're going to record our conversation with your consent. We will only utilize the recording to understand better what we can learn from your experience and how to make the prototype even better in the future.

1. Which features of the app did you like the most? Why?
2. Which features were most useful for you? How?
3. Which features of the app you did not like? Why?
4. How did your usage pattern change over time during the last 14 days? Why? What was the difference between the first and the second week?
5. Let's begin and briefly talk about the different features of the prototype:
    - Alarm setting and notification (if not already discussed before)
        - Overall impressions
        - Challenges
        - Likes and dislikes
        - Changes they recommend as per their usage
    - Usage statistics (if not already discussed before):
        - Overall impressions

  - Challenges
  - Likes and dislikes
  - Changes they recommend as per their usage
- Health data (if not already discussed before):
  - Overall impressions
  - Challenges
  - Likes and dislikes
  - Changes they recommend as per their usage
- How did the usage of the app differ between the smartphone and smartwatch?
- Is there any other feature you think we should implement in the future based on your experience?
- Did anything out of the ordinary occur during last week that would not reflect day-to-day behaviour? (e.g., travelling, vacation, stress at work or exams)
- All that's left for me to say is once more, thank you for taking part! We hope we were able to raise your awareness of your smartphone use.

APPENDIX C

Participant data

Due to the small number of participants, it's impossible to derive generalized insights thanks to highly individual behaviour and preferences.

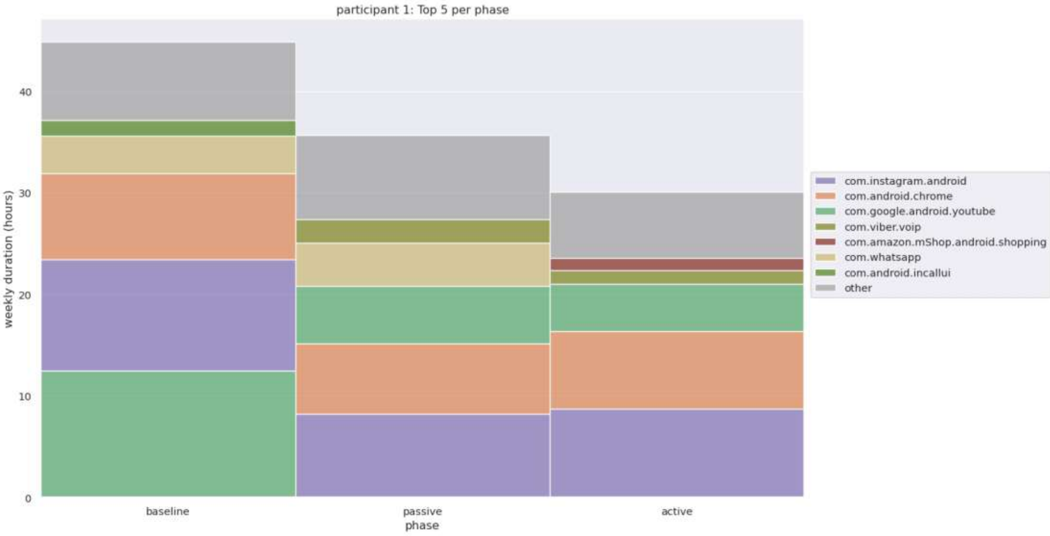


Figure C.1: Participant 1's top 5 most used applications for each week of the study

package	baseline time / rank	passive time / rank	active time / rank
com.instagram.android	10.93h / 2	8.24h / 1 **-24.60%**	8.74h / 1 **-20.04%**
com.android.chrome	8.52h / 3	6.93h / 2 **-18.68%**	7.63h / 2 **-10.51%**
com.google.android.youtube	12.49h / 1	5.65h / 3 **-54.73%**	4.66h / 3 **-62.70%**
com.viber.voip	1.47h / 6	2.28h / 5	1.32h / 4
com.amazon.mShop.android.shopping	0.08h / 25	0.33h / 14	1.22h / 5
com.android.incallui	1.60h / 5	1.08h / 6	1.18h / 6
com.whatsapp	3.65h / 4	4.26h / 4	1.00h / 7
other	6.16h ($n = 55$)	6.91h ($n = 59$) +12.27%	4.32h ($n = 48$) -29.81%
total time	**44.90h**	**35.69h**	**30.07h**
daily average	6.41h	5.10h	4.30h
total change		**-20.52%**	**-33.02%**

Table C.1: Participant 1's weekly top 5 application totals

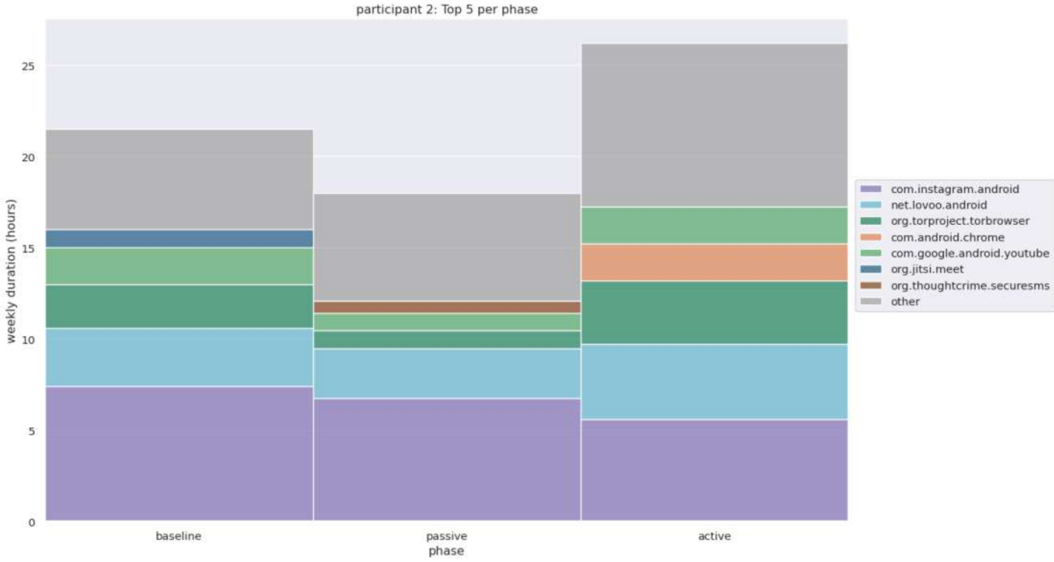


Figure C.2: Participant 2's top 5 most used applications for each week of the study

package	baseline time / rank	passive time / rank	active time / rank
com.instagram.android	7.42h / 1	6.75h / 1 **-9.07%**	5.59h / 1 **-24.69%**
net.lovoo.android	3.16h / 2	2.74h / 2 **-13.23%**	4.14h / 2 **+31.09%**
org.torproject.torbrowser	2.39h / 3	0.99h / 3	3.47h / 3
com.android.chrome	0.40h / 10	0.61h / 7	2.03h / 4
com.google.android.youtube	2.05h / 4	0.93h / 4	1.99h / 5
org.jitsi.meet	1.01h / 5	0.49h / 9	0.95h / 6
org.thoughtcrime.securesms	0.54h / 8	0.66h / 5	0.50h / 11
other	4.53h ($n = 35$)	4.83h ($n = 45$) +6.74%	7.55h ($n = 57$) +66.72%
total time	**21.50h**	**18.01h**	**26.23h**
daily average	3.07h	2.57h	3.75h
total change		**-16.20%**	**+22.02%**

Table C.2: Participant 2's weekly top 5 application totals

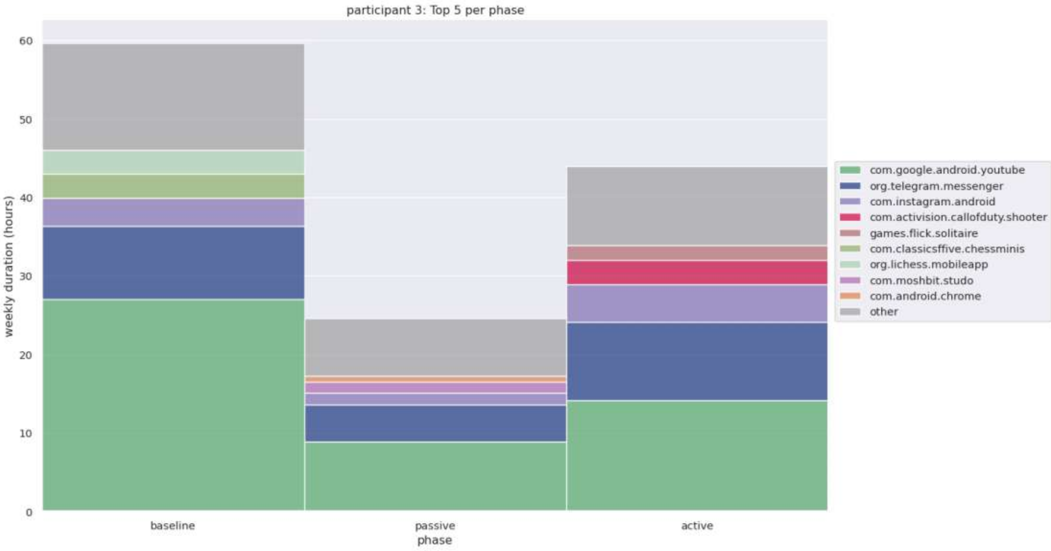


Figure C.3: Participant 3's top 5 most used applications for each week of the study

package	baseline time / rank	passive time / rank	active time / rank
com.google.android.youtube	27.00h / 1	8.89h / 1 **-67.07%**	14.13h / 1 **-47.68%**
org.telegram.messenger	9.34h / 2	4.68h / 2	9.94h / 2
com.instagram.android	3.54h / 3	1.50h / 3 **-57.66%**	4.83h / 3 **+36.52%**
com.activision.callofduty.shooter	— / —	— / —	3.07h / 4
games.flick.solitaire	— / —	— / —	1.93h / 5
com.android.chrome	2.90h / 6	0.82h / 5	1.27h / 6
com.moshbit.studo	1.27h / 7	1.39h / 4	0.79h / 9
org.lichess.mobileapp	3.01h / 5	0.12h / 24	0.37h / 13
com.classicsffive.chessminis	3.14h / 4	0.29h / 12	0.04h / 39
other	9.48h ($n = 72$)	6.87h ($n = 72$) -27.48%	7.55h ($n = 81$) -20.31%
total time	**59.68h**	**24.57h**	**43.92h**
daily average	8.53h	3.51h	6.27h
total change		**-58.84%**	**-26.41%**

Table C.3: Participant 3's weekly top 5 application totals

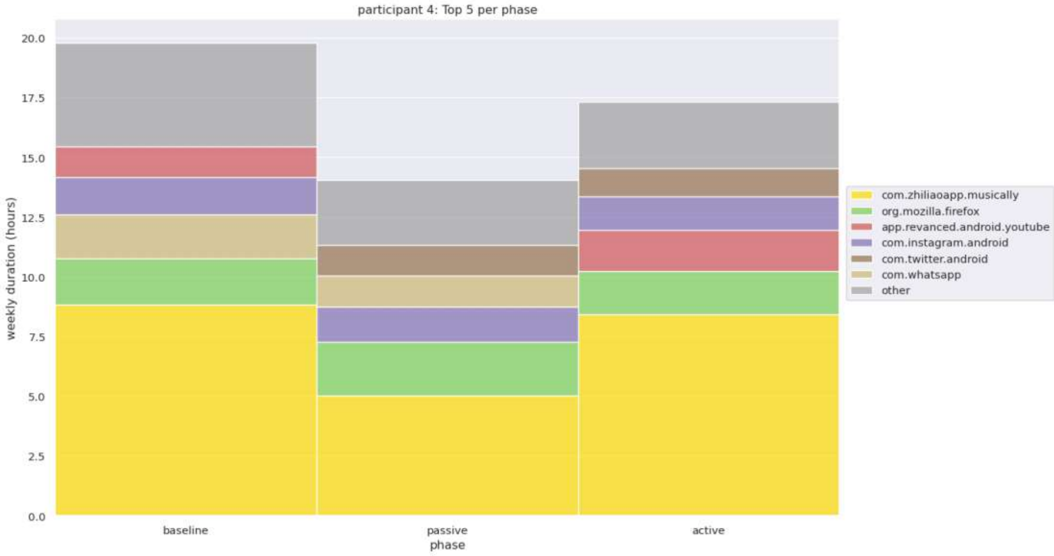


Figure C.4: Participant 4's top 5 most used applications for each week of the study

package	baseline time / rank	passive time / rank	active time / rank
com.zhiliaoapp.musically	8.83h / 1	5.04h / 1 **-42.93%**	8.43h / 1 **-4.55%**
org.mozilla.firefox	1.96h / 2	2.23h / 2	1.82h / 2
app.revanced.android.youtube	1.27h / 5	0.32h / 7	1.72h / 3
com.instagram.android	1.55h / 4	1.48h / 3	1.39h / 4
com.twitter.android	0.79h / 7	1.29h / 5	1.17h / 5
com.whatsapp	1.83h / 3	1.30h / 4	0.73h / 6
other	3.57h ($n = 38$)	2.39h ($n = 45$) -32.89%	2.06h ($n = 32$) -42.23%
total time	**19.80h**	**14.05h**	**17.33h**
daily average	2.83h	2.01h	2.48h
total change		**-29.04%**	**-12.47%**

Table C.4: Participant 4's weekly top 5 application totals

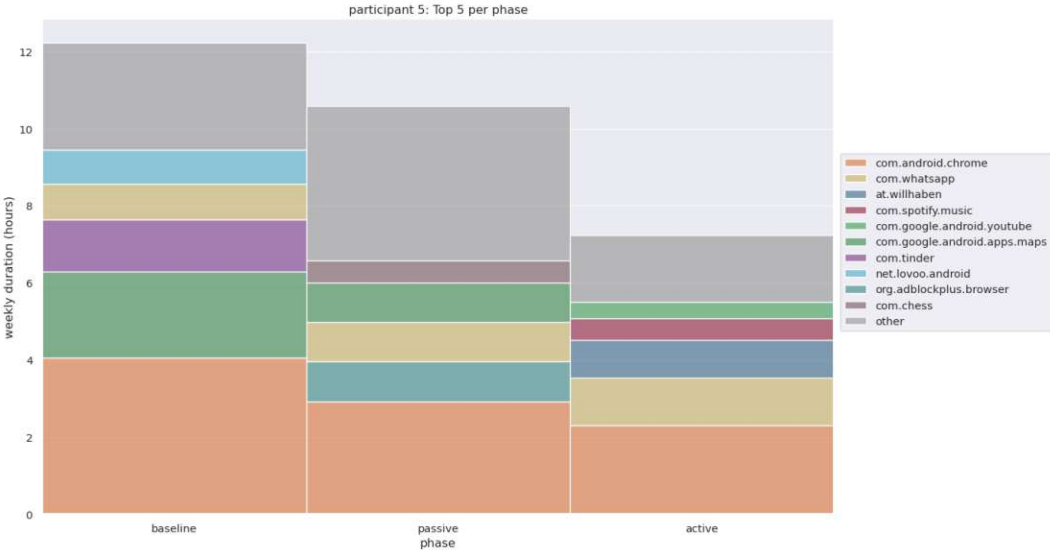


Figure C.5: Participant 5's top 5 most used applications for each week of the study

package	baseline time / rank	passive time / rank	active time / rank
com.android.chrome	4.05h / 1	2.91h / 1 **-28.11%**	2.31h / 1 **-43.00%**
com.whatsapp	0.93h / 4	1.02h / 3	1.22h / 2
at.willhaben	0.55h / 8	0.34h / 10	0.98h / 3
com.spotify.music	0.60h / 7	0.47h / 9	0.57h / 4
com.google.android.youtube	0.01h / 25	— / —	0.42h / 5
org.adblockplus.browser	0.60h / 6	1.05h / 2	0.26h / 8
com.google.android.apps.maps	2.25h / 2	1.01h / 4	0.02h / 20
com.tinder	1.34h / 3	0.52h / 7	— / —
net.lovoo.android	0.88h / 5	0.54h / 6	— / —
com.chess	— / —	0.59h / 5	— / —
other	1.12h ($n = 32$)	2.18h ($n = 37$) +94.65%	1.47h ($n = 34$) +31.41%
total time	**12.34h**	**10.65h**	**7.25h**
daily average	1.76h	1.52h	1.04h
total change		**-13.71%**	**-41.26%**

Table C.5: Participant 5's weekly top 5 application totals

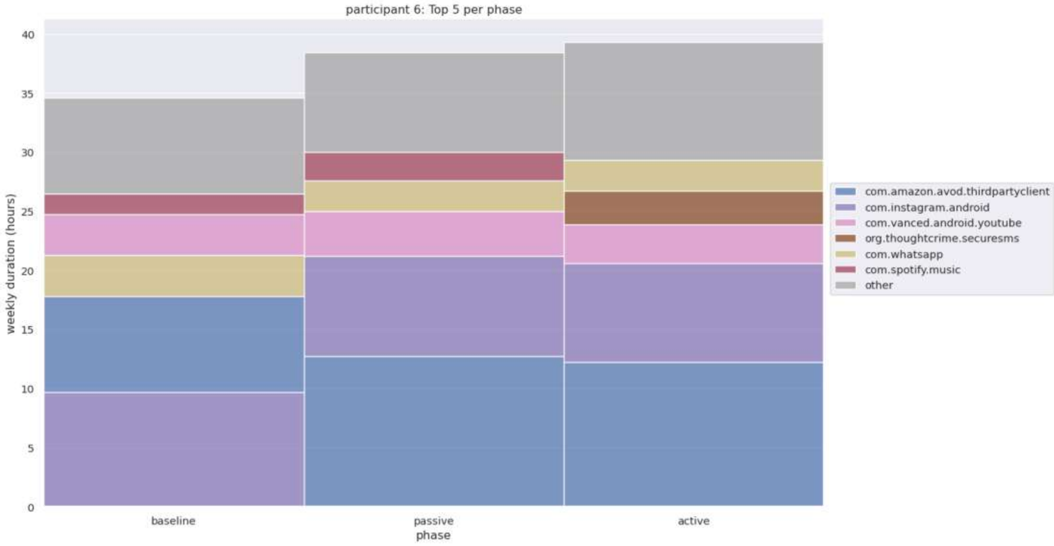


Figure C.6: Participant 6s's top 5 most used applications for each week of the study

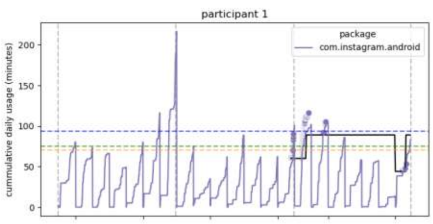
package	baseline time / rank	passive time / rank	active time / rank
com.amazon.avod.thirdpartyclient	8.12h / 2	12.75h / 1 **+57.12%**	12.25h / 1 **+50.96%**
com.instagram.android	9.68h / 1	8.44h / 2 **-12.75%**	8.37h / 2 **-13.51%**
com.vanced.android.youtube	3.45h / 4	3.81h / 3 **+10.46%**	3.27h / 3 **-5.10%**
org.thoughtcrime.securesms	1.53h / 6	2.08h / 6	2.84h / 4
com.whatsapp	3.49h / 3	2.60h / 4	2.56h / 5
com.spotify.music	1.74h / 5	2.38h / 5	1.83h / 6
other	6.62h ($n = 63$)	6.37h ($n = 69$) -3.71%	8.18h ($n = 78$) +23.65%
total time	**34.62h**	**38.43h**	**39.32h**
daily average	4.95h	5.49h	5.62h
total change		**+11.03%**	**+13.59%**

Table C.6: Participant 6's weekly top 5 application totals

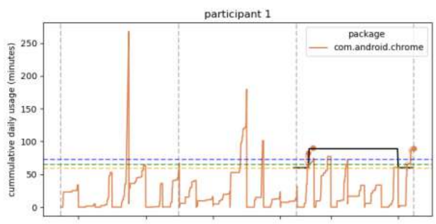
APPENDIX D

Daily usage of "problematic" applications in respect to defined limits

Figures D.1a to D.6c show daily (00:00 to 23:59) cumulative screen time of applications for which participants defined alarms. The black lines represent alarm limits at a given time, whereas points in color signify triggered alarms. Dashed horizontal lines represent daily averages per phase (*baseline*: blue, *passive*: orange, *active*: green), whereas vertical dashed lines mark phase start/end timestamps.

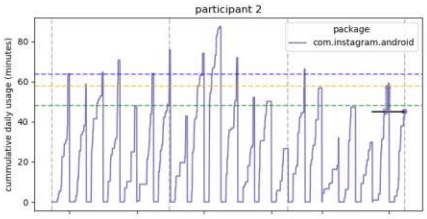


(a) *com.instagram.android*/*Instagram*

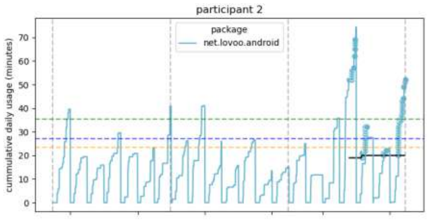


(b) *com.android.chrome*/*Google Chrome*

Figure D.1: Participant 1's usage of *com.instagram.android* and *com.android.chrome*

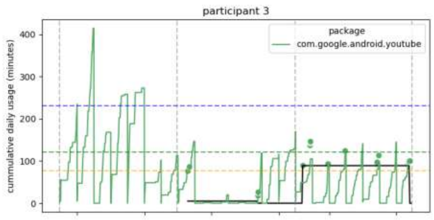


(a) *com.instagram.android/Instagram*

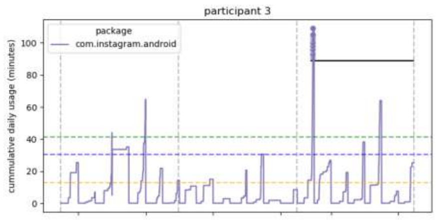


(b) *net.lovoo.android/Lovoo*

Figure D.2: Participant 2's usage of *com.instagram.android* and *net.lovoo.android*



(a) *com.google.android.youtube/Youtube*



(b) *com.instagram.android/Instagram*

Figure D.3: Participant 3's usage of *com.google.android.youtube* and *com.instagram.android*

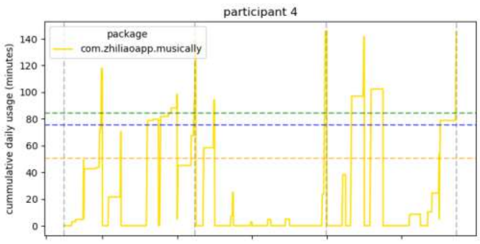


Figure D.4: Participant 4's usage of *com.zhiliaoapp.musically/TikTok*

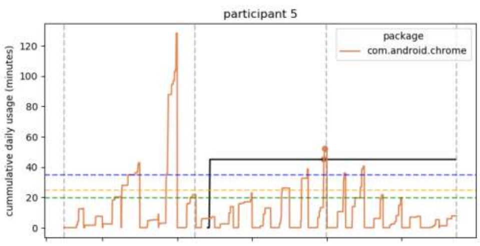
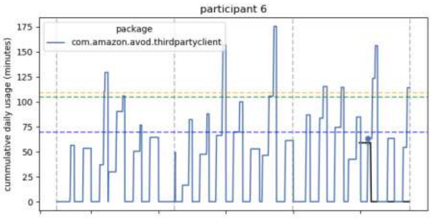
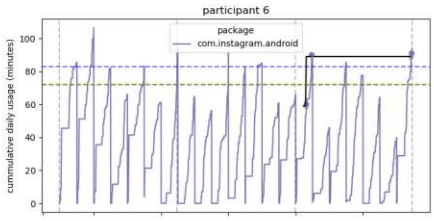


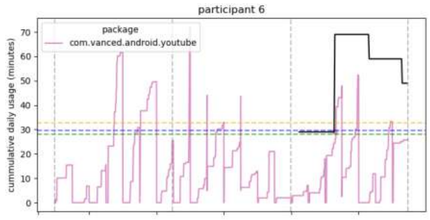
Figure D.5: Participant 5's usage of *com.android.chrome/Google Chrome*



(a) *com.amazon.avod.thirdpartyclient/ Amazon Prime*



(b) *com.instagram.android/ Instagram*



(c) *com.vanced.android.youtube Youtube Vanced*

Figure D.6: Participant 6's usage of *com.amazon.avod.thirdpartyclient*, *com.instagram.android* and *com.vanced.android.youtube*

APPENDIX E

Timeline analysis

phase	mean	change
baseline	10.92m	—
passive	8.39m	-2.53m
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = \mathbf{0.00002}, e = 0.51$	
active	9.74m	-1.18m
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = 0.0678$	
nightly baseline	10.90m	—
nightly passive	7.88m	-3.02m
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = \mathbf{0.00073}, e = 0.47$	
nightly active	8.70m	-2.20m
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = \mathbf{0.00073}, e = 0.32$	
daily baseline	10.94m	— —
daily passive	8.90m	-2.04m
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = \mathbf{0.00049}, e = 0.69$	
daily active	10.79m	-0.16m
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = 0.2593$	

Table E.1: Hourly screentime including null values

phase	mean	change
daily|nightly		
baseline	10.94m|10.90m	$0.06m^*$
day v.s. night	Mann-Whitney U:	$H_0 : F(u) = G(u)$
		$H_1 : \exists u : F(u) \neq G(u)$
	$p = 0.5899$	
passive	8.90m|7.88m	$1.05m^*$
day v.s. night	Mann-Whitney U:	$H_0 : F(u) = G(u)$
		$H_1 : \exists u : F(u) \neq G(u)$
	$p = 0.9323$	
active	10.79m|8.70m	$2.09m^*$
day v.s. night	Mann-Whitney U:	$H_0 : F(u) = G(u)$
		$H_1 : \exists u : F(u) \neq G(u)$
	$p = 0.4776$	

Table E.2: Hourly screentime including null values

phase	mean	change
baseline	16.51m	—
passive	12.93m	-3.58m
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = \mathbf{0.00326}, e = 0.27$	
active	14.96m	-1.55m
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = 0.0938$	
nightly baseline	18.80m	—
nightly passive	15.60m	-3.20m
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = 0.1018$	
nightly active	16.45m	-2.34m
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = \mathbf{0.00171}, e = 0.39$	
daily baseline	14.72m	— —
daily passive	11.22m	-3.50m
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = \mathbf{0.00024}, e = 1.45$	
daily active	13.94m	-0.78m
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = 0.3110$	

Table E.3: Hourly screentime cleaned from null values

phase	mean	change
daily|nightly		
baseline	14.72m|18.80m	$-2.12m^*$
day v.s. night	Mann-Whitney U:	$H_0 : F(u) = G(u)$
		$H_1 : \exists u : F(u) \neq G(u)$
	$p = 0.0519$	
passive	11.22m|15.60m	$-5.61m^*$
day v.s. night	Mann-Whitney U:	$H_0 : F(u) = G(u)$
		$H_1 : \exists u : F(u) \neq G(u)$
	$p = 0.0887$	
active	13.94m|16.45m	$0.11m^*$
day v.s. night	Mann-Whitney U:	$H_0 : F(u) = G(u)$
		$H_1 : \exists u : F(u) \neq G(u)$
	$p = 0.7125$	

Table E.4: Hourly screentime cleaned from null values

phase	mean	change		
baseline	13.15m	—		
passive	10.45m	-2.70m		
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$		
		$H_0 : F(d)$	symmetric about	$\mu = 0$
		$H_1 : F(d)$	symmetric about	$\mu > 0$
	$p = \mathbf{0.02821}, e = 0.15$			
active	10.81m	-2.33m		
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$		
		$H_0 : F(d)$	symmetric about	$\mu = 0$
		$H_1 : F(d)$	symmetric about	$\mu > 0$
	$p = \mathbf{0.02284}, e = 0.37$			
nightly baseline	15.76m	—		
nightly passive	13.56m	-2.19m		
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$		
		$H_0 : F(d)$	symmetric about	$\mu = 0$
		$H_1 : F(d)$	symmetric about	$\mu > 0$
	$p = 0.3955$			
nightly active	11.35m	-4.41m		
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$		
		$H_0 : F(d)$	symmetric about	$\mu = 0$
		$H_1 : F(d)$	symmetric about	$\mu > 0$
	$p = \mathbf{0.00024}, e = 0.58$			
daily baseline	10.87m	—		—
daily passive	8.17m	-2.71m		
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$		
		$H_0 : F(d)$	symmetric about	$\mu = 0$
		$H_1 : F(d)$	symmetric about	$\mu > 0$
	$p = \mathbf{0.00024}, e = 0.86$			
daily active	10.43m	-0.44m		
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$		
		$H_0 : F(d)$	symmetric about	$\mu = 0$
		$H_1 : F(d)$	symmetric about	$\mu > 0$
	$p = 0.4250$			

Table E.5: Hourly screentime of “limited” applications including null values

phase	mean	change
daily|nightly		
baseline	10.87m|15.76m	$-3.55m^*$
day v.s. night	Mann-Whitney U:	$H_0 : F(u) = G(u)$
		$H_1 : \exists u : F(u) \neq G(u)$
	$p = 0.1978$	
passive	8.17m|13.56m	$-7.05m^*$
day v.s. night	Mann-Whitney U:	$H_0 : F(u) = G(u)$
		$H_1 : \exists u : F(u) \neq G(u)$
	$p = \mathbf{0.00291}, e = 0.96$	
active	10.43m|11.35m	$0.20m^*$
day v.s. night	Mann-Whitney U:	$H_0 : F(u) = G(u)$
		$H_1 : \exists u : F(u) \neq G(u)$
	$p = 0.7125$	

Table E.6: Hourly screentime of “limited” applications including null values

phase	mean	change
baseline	26.51%	—
passive	23.83%	-2.68%
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = 0.0758$	
active	26.17%	-0.34%
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = 0.5168$	
nightly baseline	26.49%	—
nightly passive	21.84%	-4.65%
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = \mathbf{0.00049}, e = 0.33$	
nightly active	21.98%	-4.51%
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = \mathbf{0.00049}, e = 0.33$	
daily baseline	26.53%	— —
daily passive	25.83%	-0.70%
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = 0.2593$	
daily active	30.38%	3.85%
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = 0.9998$	

Table E.7: Hourly proportion of “limited” apps’screen time of total screentime including null values

phase	mean	change
daily|nightly		
baseline	26.53%|26.49%	0.10%*
day v.s. night	Mann-Whitney U:	$H_0 : F(u) = G(u)$
		$H_1 : \exists u : F(u) \neq G(u)$
	$p = 0.5512$	
passive	25.83%|21.84%	4.08%*
day v.s. night	Mann-Whitney U:	$H_0 : F(u) = G(u)$
		$H_1 : \exists u : F(u) \neq G(u)$
	$p = 0.9323$	
active	30.38%|21.98%	8.46%*
day v.s. night	Mann-Whitney U:	$H_0 : F(u) = G(u)$
		$H_1 : \exists u : F(u) \neq G(u)$
	$p = 0.1600$	

Table E.8: Hourly proportion of "limited" apps'screen time of total screentime including null values

phase	mean	change
baseline	57.35%	—
passive	56.95%	-0.41%
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = 0.8347$	
active	56.53%	-0.83%
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = 0.3317$	
nightly baseline	61.77%	—
nightly passive	61.85%	0.08%
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = 0.9739$	
nightly active	57.14%	-4.63%
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = \mathbf{0.00464}, e = 0.52$	
daily baseline	53.51%	— —
daily passive	53.35%	-0.16%
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = 0.5452$	
daily active	56.09%	2.58%
v.s. baseline	Wilcoxon signed-rank:	let $d = X_i - Y_i$
		$H_0 : F(d)$ symmetric about $\mu = 0$
		$H_1 : F(d)$ symmetric about $\mu > 0$
	$p = 0.9995$	

Table E.9: Hourly proportion of “limited” apps’screen time of total screentime cleaned from null values

phase	mean	change
daily|nightly		
baseline	53.51%|61.77%	$-13.99\%^*$
day v.s. night	Mann-Whitney U:	$H_0 : F(u) = G(u)$
		$H_1 : \exists u : F(u) \neq G(u)$
	$p = \mathbf{0.02842}, e = 1.07$	
passive	53.35%|61.85%	$-18.47\%^*$
day v.s. night	Mann-Whitney U:	$H_0 : F(u) = G(u)$
		$H_1 : \exists u : F(u) \neq G(u)$
	$p = \mathbf{0.00291}, e = 1.36$	
active	56.09%|57.14%	$-2.76\%^*$
day v.s. night	Mann-Whitney U:	$H_0 : F(u) = G(u)$
		$H_1 : \exists u : F(u) \neq G(u)$
	$p = 0.6707$	

Table E.10: Hourly proportion of "limited" apps'screen time of total screentime cleaned from null values