Flower: A Visualization Tool for Reflecting on Smartphone Use

Shifting Focus from Time to Context and Purpose through Self-Categorization

Master's thesis in Computer science and engineering

STINA HANSSON HANNA ADENHOLM

Department of Computer Science and Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Gothenburg, Sweden 2025

Master's thesis 2025

Flower: A Visualization Tool for Reflecting on Smartphone Use

Shifting Focus from Time to Context and Purpose through Self-Categorization

STINA HANSSON
HANNA ADENHOLM

UNIVERSITY OF
GOTHENBURG

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Computer Science and Engineering
Chalmers University of Technology
University of Gothenburg
Gothenburg, Sweden 2025

Flower: A Visualization Tool for Reflecting on Smartphone Use
Shifting Focus from Time to Context and Purpose through Self-Categorization
STINA HANSSON HANNA ADENHOLM

Supervisor: Alexandra Weilenmann, Department of Applied Information Technology
Examiner: Morten Fjeld, Department of Computer Science and Engineering

Master's Thesis 2025
Department of Computer Science and Engineering
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Typeset in LaTeX
Gothenburg, Sweden 2025

Flower: A Visualization Tool for Reflecting on Smartphone Use
Shifting Focus from Time to Context and Purpose through Self-Categorization
HANNA ADENHOLM, STINA HANSSON
Department of Computer Science and Engineering
Chalmers University of Technology and University of Gothenburg

Abstract

This thesis explores the design considerations for creating a tool that encourages meaningful reflection on smartphone use. To investigate this, an app called Flower was developed, allowing users to categorize each app session and then view visualizations of their usage. Two visualization styles were tested; a conventional bar chart and a more novel "flower field" where each flower represented an app session. Over an eight-day user study with twelve university students, findings indicated that the act of self-categorization prompted greater reflection and that users felt the visualized data better represented their experiences. While most participants preferred the flower field for its aesthetics and exploratory nature, some valued the bar chart for its clarity and ease of interpretation.

The study suggests that encouraging users to reflect on their smartphone use in context, rather than labeling all screen time as negative can motivate users to explore their data, and draw more meaningful insights about their phone use. Novel visualizations can increase user interest, though care must be taken to balance abstractness with clarity. Limitations of the study include minor technical issues and a small, homogeneous participant group. Future research could explore more personalized visualizations, longer-term use, and broader participant samples.

This work contributes to understanding how creative, user-driven tools can foster reflection and more nuanced perceptions of digital well-being.

Keywords: information visualization, reflection, screen time, smartphone use, digital well-being, interaction design

Acknowledgements

We would like to express our heartfelt gratitude to our supervisor, Alexandra Weilenmann, for her invaluable guidance, insightful feedback, and continuous support throughout the entire project. We are also deeply thankful to our family and friends for their encouragement and unwavering belief in us, which made this work possible. Finally, a sincere thank you to everyone who participated in our study, this project would not have been possible without your contributions.

Hanna Adenholm and Stina Hansson, Gothenburg, 2025-06-10

Contents

List of Figures

List of Tables

1

Introduction

The term screen time is one often heard today, and media's reporting has often painted a quite dire picture, leading to stress and guilt among people [1]. Smartphone use has been compared to gambling and drug addiction [2], and a lot of research has been done to try to determine the amount of screen time that leads to negative effects on users' mental and physical health [3]. While some connections and conclusions can be made, results are still inconclusive and sometimes contradictory [3], [4].

In recent years, screen time has also become a central topic in public health discussions, policy debates, and parenting discourse [4], [5]. Government agencies and public health authorities in various countries have begun issuing screen time guidelines, often recommending a maximum number of hours per day, especially for children and adolescents [4]. While these guidelines are well-intentioned, they tend to take a one-size-fits-all approach that emphasizes total duration over context, purpose, or type of screen use. This risks oversimplifying the issue and promoting a view that all screen time is equally harmful, regardless of whether it involves passive scrolling or active learning, connection, or creativity. As a result of this, the focus of previous research has been on minimizing the amount of time spent on screen use [3].

Additionally, the term screen time is often used to mean different things, and Kaye, Orben, Ellis, *et al.* [6] see a lack of standardizations and variations in terminology between researchers. There are definitions of screen time that includes all time spent in front of any screen while other definitions are less inclusive, only including for example, time spent passively watching screen-based entertainment [6].

While previous research has mostly revolved around the quantity of screen time and its impact, it is moving towards a more nuanced view on digital media [7]. Lanette and Mazmanian [2] argue that the common description of smartphone use as an addiction is harmful, and with smartphones being so multifaceted, researchers need to address this complexity. They suggest that it is better to encourage curiousness and reflection about smartphone use rather than guilt. Researchers such as Livingstone [8] and Orben [3] argue for moving away from screen time, since the focus on a specific time limit leave out the importance of the content and context.

So far, research revolving tools for smartphone use has mainly focused on interventions and monitoring use by measuring usage time [9], and little attention has been placed on the way smartphone use is displayed to the user, how it is visualized and

what categories it is divided into. Sosa-Tzec [10] proposes being more innovative and less conventional when creating visualizations of screen time, as it may make the user more interested in reflecting about their screen time.

While not much research has been done on visualizations of screen time, similar work exist in broader fields, such as personal visualizations. Personal visualizations are visualizations meant for a personal context, for non-experts in their daily life, in contrast to information visualization, which traditionally involves designing visualization for experts to use in their work [11]. For instance, researchers have explored how people respond to abstract versus conventional visualizations when applied to data, such as daily water intake [12] and social media activity on Facebook[13].

In our thesis, we want to explore how smartphone use can be visualized in another way than just an amount of hours on the screen. Are there ways of visualizing the information that instead puts an emphasis on meaningful smartphone use and reflection?

1.1 Aim

This thesis aims to investigate how an application with a focus on visualizing smartphone use that represents the user's behavior can be designed and implemented. Rather than putting emphasis on the total amount of time, this application would be designed around reflection and letting the users themselves evaluate how they feel about their smartphone habits. The areas of focus are both what impact self-categorizing smartphone use has, but also what impact it has to also visualize this data back to the user.

1.2 Research questions

Based on the aim, a research question and two sub-questions are formulated.

1. What design considerations are important when creating a digital tool that encourages reflection on smartphone use, with a focus on promoting meaningful rather than minimized usage?
   (a) What are the effects of letting the user themselves evaluate their smartphone use depending on context and purpose?
   (b) How does visualizing users evaluation of their smartphone use influence their reflection and behavior?

1.3 Delimitations

Due to the limited time of this project, some delimitations are made in order to narrow the scope. Firstly, only young adults around the age of 18-30 are considered. This is because the young adult population is more prone to problematic smartphone use [14], and it was also the age group most accessible for us to reach out to in

order to to recruit participants for the study. Secondly, only smartphones and no other devices with screens, such as tablets or laptops are considered. The reason for this is mainly to ease the gathering of data, as not multiple versions of the application for different devices has to be created. Furthermore, only participants with Android devices are included in the user testing of the application, otherwise two separate apps would have to be created, which was not feasible in this time frame. Additionally, the available tools were not compatible with development for iOS.

2

Theory

In this chapter, theoretical frameworks and concepts that underpin the thesis are presented. It begins by exploring various definitions and models of digital well-being, a core concept in this thesis. Following this, another important aspect of this project is brought up, information visualization. Then, time as a concept is explored more deeply in order to give a solid foundation for the thesis. Finally, different possible ways of categorizing smartphone use are presented, including distinctions such as active vs. passive use, instrumental vs. habitual, and regretful use.

2.1 Digital well-being

One common term that is often used when talking about screen time is digital well-being. Digital well-being is a new term within the field of Human-Computer Interaction (HCI) research, and is becoming a common term to describe the way technologies affect people's health, both mental, physical, social, and emotional [15]. Since it is still a quite new term, it does not have one clear definition, and can therefore be used differently by different researchers. A distinction can be made between the technology that is designed to improve well-being, and the considerations made to reduce the negative effects on users, which is sometimes called positive computing [15]. Calvo and Peters [16] argues that well-being often has been overlooked when designing new technologies, but observes an ongoing shift toward positive computing, which they define as "the design and development of technology to support psychological well-being and human potential". To support this shift, Calvo and Peters [16] explains that multidisciplinary collaboration is necessary, since one field alone does not have the needed tools and methods.

2.1.1 A dynamic model of digital well-being

Vanden Abeele [17] uses the definition of digital well-being as a "subjective individual experience of optimal balance between the benefits and drawbacks obtained from mobile connectivity", and proposes a dynamic model of digital well-being. This model explores the connection between person-specific, device-specific and context-specific factors. Person-specific factors can include different personality traits that a person can have that makes them more likely to have problems with digital media. It can also include how one person's internal state, for example their mood, can have an impact. Device-specific factors refers to the way devices and apps are designed,

often in order to keep users engaged, which can make them hard to resist. Context-specific factors goes into how the context can dictate how people use their phones. Some contexts, like movie theaters or meetings, clearly calls for disconnection, while other contexts can be more ambiguous. Vanden Abeele [17] explains that this model can be used to better accommodate the complexity of digital well-being, one that not only defines digital well-being as the absence of bad habits, but rather the balance between connectivity and disconnectivity.

2.1.2 Hedonistic and eudaimonic experiences

Lukoff, Yu, Kientz, *et al.* [18] compare the hedonistic tradition of well-being, the view that happiness is the presence of positive experiences and absence of negative ones, to the eudaimonic tradition, where well-being is is based on a sense of fulfillment and meaning. Applied to smartphone use, scrolling through pictures of cute animals can be a hedonistic experience, while messaging someone to apologize might be an eudaimonic experience [18]. Vanden Abeele [17] has a view of digital well-being where both hedonic and eudaimonic experiences are considered, but acknowledges that it is hedonic experiences that makes it hard for people to resist their phone.

2.1.3 Encouraging mindfulness and awareness

One factor that has been consistently linked to increased well-being is mindfulness [19]. Over the past decade, the HCI field has shown growing interest in mindfulness as a way to promote healthier relationships with technology [20], [21]. While definitions of mindfulness vary, it generally involves awareness of both internal and external experiences and their effects [19]. It involves being conscious of ones actions and their impact, particularly in everyday routines and decision-making [20]. Related concepts include reflection, self-monitoring and self-awareness [19], [21].

In the context of smartphone use, mindfulness can encourage users to actively reflect on when, why and how they use their smartphones. This raised awareness can help users to recognize certain triggers that can lead to obsessive phone use. Common triggers include downtime, boring tasks, social awkwardness and feelings of anticipation [22]. By making triggers such as these visible, it can help users distinguish between intentional and unintentional use, and help them to make more deliberate choices. They can learn to recognize when they use an app with no specific goal or meaning and then choose to disengage from it [21]. As such, reflection could be a key mechanism for disrupting automatic, often meaningless phone use, and fostering a more conscious and purposeful use.

2.1.4 Behavioral change and digital well-being

Self determination theory, or SDT, is a framework for studying motivation, personality development and well-being [23]. Three basic psychological needs are described; autonomy, relatedness and competence. If these needs are supported, the well-being of a person is said to be higher, while the opposite is true if these needs are hindered. Peters, Calvo, and Ryan [24] propose the METUX model, a model based on SDT

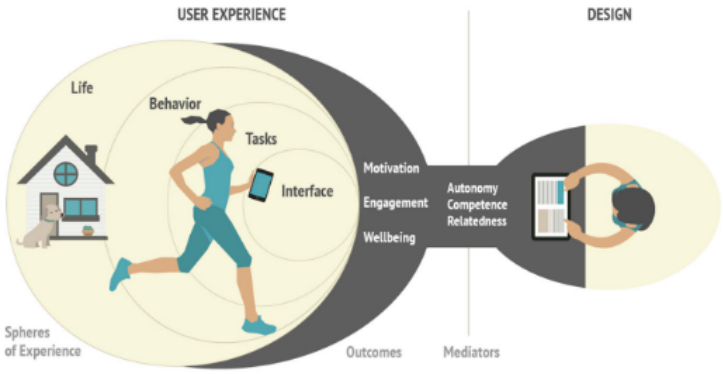


Figure 2.1: The METUX model. Source: [24]

applied on digital well-being, to help designers and researchers create technology that supports these needs. In this context, the autonomy need refers to individuals acting with a strong sense of willingness and alignment with their personal goals and values. The second need, competence, is about the individual feeling capable and productive. Thirdly, relatedness gives a sense of belonging and feeling connected to others. The METUX model, see figure 2.1, has six "spheres of experience", where these needs should be considered. The first, adoption, is how and why a user acquire the new technology, and what they expect of it. The next sphere is the interface itself and how the experience of using it is. Further, the task sphere is the task that the interface enables the user to do, and the extent to which it helps the user perform it. The next sphere is the overarching behavior that the technology support, and how successfully it does this. Then, the life sphere is the impact of the technology on the user's overall life and their well-being. Even larger, the last sphere is society, that goes beyond the user and evaluates if the technology has any impact on the society surrounding the user.

2.2 Information visualization

The field of information visualization is vast, but in broad terms it can be described as the creation of graphical representations of information, typically generated with a computer [25]. The data is often abstract and non-spatial, making it one very important part of information visualization to present the data in a way so that it is intuitive and meaningful. Positioned on the intersection of art and science, it is both a creative process where the designer communicates a message and inspires users, but it must also be done in an accurate and rigorous way [25]. Chen [25] describes the ultimate goal of information visualization as helping users gain insights, such as achieving a deeper understanding or experiencing breakthroughs, and calls it the holy grail of the field.

Moere and Purchase [26] relate a model for design requirements within architecture, the Vitruvius triangle, to information visualization, see figure 2.2. The three requirements are utility, soundness, and attractiveness. Utility has to do with functionality and usefulness, and can often be quantitatively measured by how effective and ef-

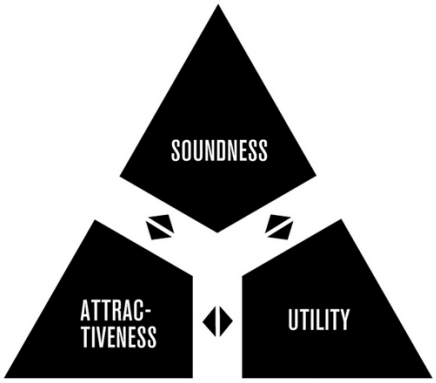


Figure 2.2: Vitruvius triangle, a model from architecture applied to information visualization. Source: [26]

ficient the visualization is. The second requirement, soundness, instead focuses on how robust the visualization is. For example, how well the visualization can be used for other data sets, how resilient it is to differently complex data and so forth. Thirdly, the attractiveness requirement goes into the aesthetics of the visualization, how appealing and beautiful it is, or the novelty of it. While it is common for academic research to explain the methodology behind the utility and soundness requirements, there is no expectation of this for the attractiveness requirement, which means there has been less advancement in this aspect of the field [26]. However, there is research showing that aesthetics can support the utility of a visualization, motivate users to engage with the visualization, and make them remember it better afterwards [26].

2.2.1 Metaphors

Metaphors, originally a term from linguistics, has been used extensively within information visualization. Cox [27] coined the term "visaphor", a visual metaphor, and how they can be used to apply the understanding from one domain onto another. A metaphor can affect how we think about a certain topic, for example "time is money", brings connotations from how we view money, and apply them to how we view time. However, Cox [27] describes this process as "partial and nonarbitrary", only some characteristics are mapped onto the target, which allows us to see it in a new light. Cox [27] puts emphasis on the fact that visaphors are approximations, not reality, and that visaphors are biased from those who create them, and that there is always alternative ways of representing the data.

2.3 Time as a concept

One reason as to why visualizing screen use with time could be problematic is that time can be perceived very differently from individual to individual. It can depend a lot on our cultural, historical or religious context [28]. Hoffman [28] describes how they perceived time differently when living in different cultures. How in eastern Europe, when they grew up, time was frozen, people weren't in a hurry to get

anywhere, they did not worry about returns or results and valued long discussion late into the night. How in America, they instead perceived time as being money, they had a sort of constant worry about not having enough time, that time was very valuable. In Britain they instead perceived people as being more decisive and unhurried.

However, Hoffman [28] argues that the perception of time are changing, the pre-modern rhythms and habits that existed are now becoming more productive but ruthless. With the forthcoming of computer clocks, of communications and the connectivity today, one can exist in every single time zone all at once. One can enjoy summer and winter several times per year by traveling. With this shift, time is becoming more and more deterritorialized and virtual [28]. It is becoming more universal.

Another aspect of time is how it can be perceived in the moment or more short term. There exists some research on how time distortion, dissociation between actual and perceived time intervals, can come as an effect of excessive phone use [29]. Two different types of time distortion has been identified in correlation to compulsive technology use. The first is that the perceived length of usage sessions is shorter than they really are, motivating the user to continue. The other is that the perceived length of a non-digital short task is perceived as longer [29].

2.4 Ways of categorizing smartphone use

In the body of work, there are multiple ways to think about and to divide smartphone use into different taxonomies. Some are more objective, while some are up to the user to define for themselves. In order to give a clear understanding and overview of these terms, some of these distinctions are presented in the following section.

2.4.1 Active and passive use

Social media use can be divided into active and passive use, where active use is the time the user spends actively interacting with others on the app, chatting and commenting, or posting something of their own. In contrast, passive use is when the user only consumes what others has already posted, and not interacting with the content [30]. An extension of the model divides active and passive use into further categories [30]. Active use can include both targeted communication; for one specific person or group, and non-targeted communication; public and for no specific target. Active use can also both be warm or cold, depending on if the communication is positive or negative. Similarly, the effect of passive use can also vary depending on how relevant the content is to the person viewing it. Negative self-comparison can increase if viewing content of other similar people, rather than content that is more irrelevant to the person. Similarly, viewing other peoples' achievement can lead to damaging comparison, rather than viewing peoples' failures. Studies have shown that passive social media use is negatively correlated with digital well-being, while active social media use was found to have a more positive impact on the digital well-being [31], but this should be viewed through the lens of the active-passive model

having been updated since.

2.4.2 Instrumental and habitual use

Lukoff, Yu, Kientz, *et al.* [18] apply a Uses and Gratifications perspective on smartphone use, and makes the distinction of instrumental and habitual use. For instrumental use, the user intentionally uses a technology in order to achieve something specific. Habitual use on the other hand, is when users only seek to pass time, and do not have a specific goal in mind. This can also include compulsive use, where users pick up their phone and go into certain apps out of habit, seeking rewards [32].

2.4.3 General and absentminded smartphone use

Terzimehi, Haliburton, Greiner, *et al.* [21] instead use the term absentminded smartphone use to describe phone use with no specific purpose, like compulsive checking and pointless scrolling. General use, is then the smartphone use that has a purpose.

2.4.4 Maladaptive mobile phone use

Maladaptive mobile phone use (MMPU) is often used to refer to excessive smartphone use that can negatively interfere with an individual's personal or professional life [33], [34]. MMPU is multi-faceted and does not have a clear consensus of its definition. Rahmillah, Tariq, King, *et al.* [34] consider the term to be an umbrella term for phenomena such as problematic smartphone use, smartphone addiction, Fear of Missing Out (FoMO), smartphone involvement, smartphone dependency, habitual phone use, possession attachment to the phone, compulsive smartphone checking and texting dependency.

While there is no single agreed-upon definition for the term, it is still important for health research due to its potentially harmful consequences [34]. For example, MMPU could lead to being distracted in traffic, increase anxiety or stress, decrease academic performance, and lead to irregular sleep patterns [33].

2.4.5 Regretful smartphone use

Another way to look at smartphone use is to view it as being either regretful or not regretful. Both Cho, Choi, Kim, *et al.* [35] and Orzikulova, Cho, Chung, *et al.* [36] used regretfulness as a framework when evaluating smartphone use. According to regret theory, regret is a counterfactual feeling that the past might have unfolded differently, especially if a different decision was made [35]. Basically, if possible rewards of alternative actions exceed the rewards from use of social media, the individual would regret the use [35].

2.4.6 A nuanced way to look at screen time

As a suggestion on how the research focus should be moving forward, Orben [3] uses an analogy of a digital diet in order to explain a more nuanced way to think about

smartphone use rather than only measuring the time. Similarly to how the type and amount of different types of food impacts the well-being of a person, the type of content and amount of time spent, has an impact. Orben [3] further argues that similarly to food, having a balanced digital diet is important, and that what is best for one person is not necessarily best for the other.

3

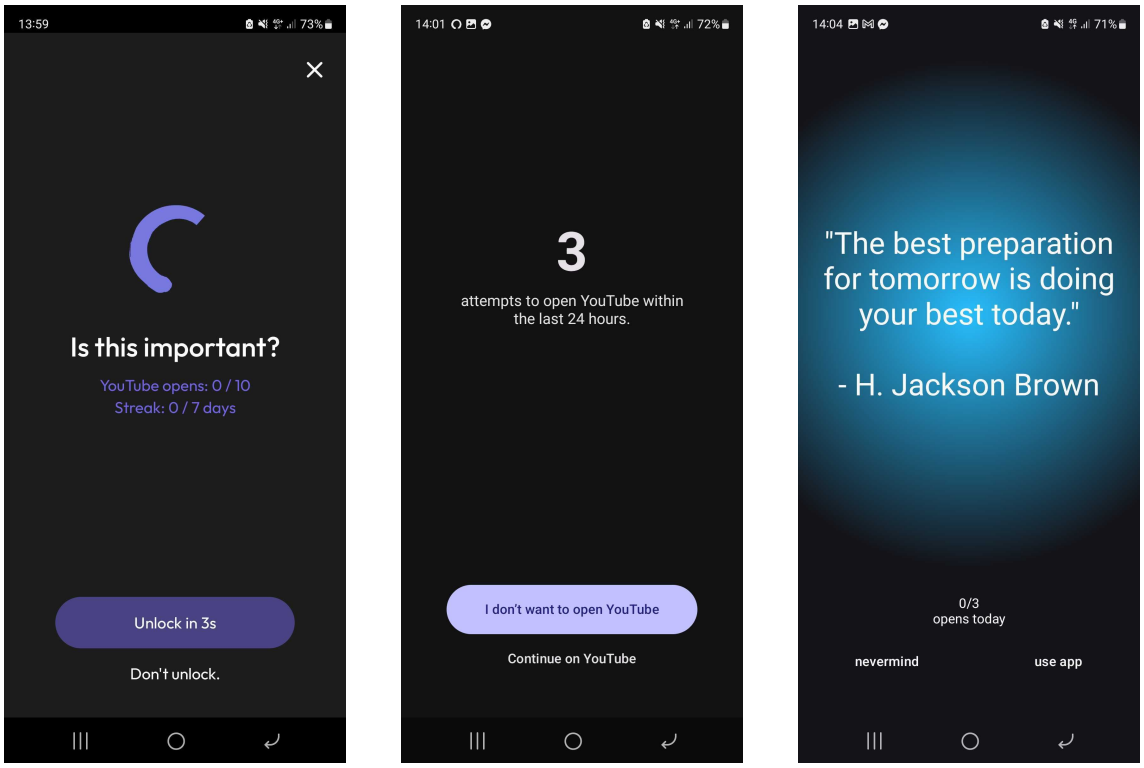
Related work

Significant research has already been conducted on the topic of smartphone use. In addition to the hundreds of applications that have been developed to assist users in regulating their screen time, numerous tools have been created for various studies. In this chapter, we will review some of these existing solutions, examining what is currently available on the market, what has already been explored, and what gaps remain. In addition, we will analyze their findings, results, and effectiveness. Understanding the market and prior research is essential, as it provides valuable insights and can inspire new ideas. Finally, the chapter discusses the limitations and shortcomings of previous work, highlighting research gaps that this study aims to address.

3.1 Tools on the market

Today, there exist hundreds of apps aiming to manage and reduce the amount of time users spend on their phones. They have however faced a lot of criticism for not effectively changing habits [33] or for not considering the different nuances of smartphone use. Apps which uses restrictive interventions to entirely lock out access to smartphones or apps can be impractical due to the dependence on phones to get through daily life. It is also impractical due to the social expectations of being constantly reachable online in relationships and in the workplace [35]. This can lead to users removing the restrictions or bypassing them, making them ineffective in affecting the users habits. Tran, Yang, Davis, *et al.* [32] suggest designing tools that move beyond the lock-out feature, since it only targets compulsive smartphone use, failing to understand the nuance of smartphone use. Other non-restrictive methods can be to provide feedback and visualization of one's phone usage. However, they seldom consider how the app is being used or the context in which it is used, making them too high-level to gain any real insights [35]. This lack of consideration for the diverse usage contexts of apps create frustration in users [35].

Rahmillah, Tariq, King, *et al.* [33] evaluate the effectiveness of such apps designed to reduce smartphone use, and pinpoints different features and evaluates their effectivity. They found 7 different types of features; self-tracking, social tracking, goal setting, blocking, gamification, simplification and assessment. Out of these, self-tracking and goal setting were most frequently used, whereas gamification and assessment were the least used [33]. They found that specifically iOS screen time



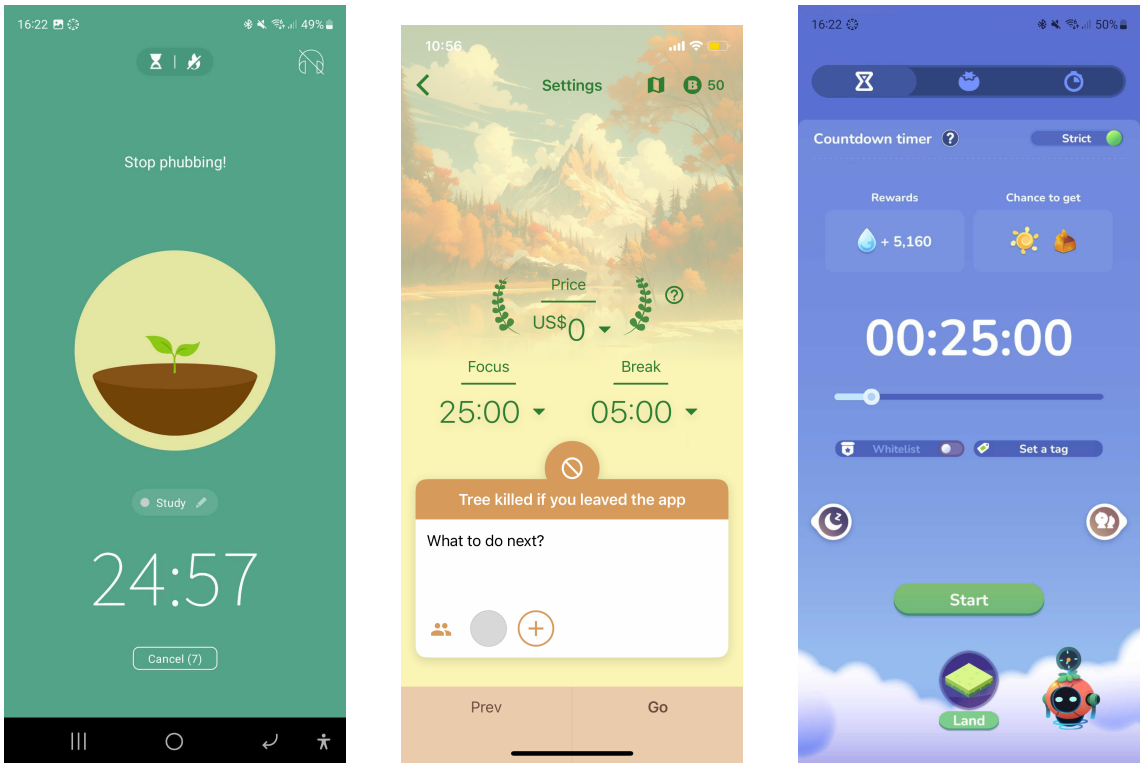
(a) ScreenZen Source: [41] (b) one sec Source: [44] (c) clearspace Source: [43]

Figure 3.1: Views of three different apps after opening a restricted app.

[37], Forest [38], Screen time [39], and AntiSocial [40] were effective in reducing screen time. Other methods for reducing screen time such as using gray-scale mode when a certain time limit is reached, having an app limit feature and using mixed interventions were also found effective. Despite these results the authors of the paper argue that future research should focus on encouraging more meaningful phone use rather than only exploring ways to reduce smartphone use. They encourage app developers to develop apps that are more "enjoyable to use, meets their needs, supports existing habits, allows for goal creation and modification and provides rewards and opportunities to share progress within a social community" [33].

The apps ScreenZen [41], one sec [42], clearspace [43] all have a similar feature of making users wait a few seconds before being able to open selected apps. After this time has passed, the user receives the question if they want to use the app, or if they want to close it, see figure 3.1. If the user decides they still want to use the app, these apps intervenes again after some time, in order to avoid that the user uses the app for longer than intended. one sec has been proven to reduce the consumption of certain apps [44], although it is worth noting that one of the authors of the paper is also the founder of one sec. Both one sec and ScreenZen also has the option to block apps, with ScreenZen also having features like streaks, and the option to block a specific feature of some apps, like Youtube Shorts or Instagram Reels.

Another app that also allows blocking specific features of some apps is StayFree [45]. It takes a more conventional approach of blocking apps after a certain amount of



(a) Forest Source: [38] (b) Flora Source: [46] (c) FocusPlant Source: [47]

Figure 3.2: Views of three different apps for focusing.

time, but also has other features, for example allowing users to compare their screen time to the global average and be able to connect multiple kinds of devices to one account.

A category of smartphone use tools are apps that does not track users' overall smartphone use, but instead, the purpose is to stay focused for a certain amount of time, and if the user don't touch their phone they get a reward. Forest [38], Flora [46] and Focus Plant [47] are examples of this, see figure 3.2. In Forest and Flora, users set a timer, and during that time the app will grow a virtual tree. If the user terminates the session by opening other apps, the tree will die. In both apps, the users also have the option to be able to indirectly plant real trees, with Flora also allowing users to set an amount of money that will go to planting real trees if the user fails to focus. Focus Plant has a similar functionality of a timer, but has a more gamified approach. When the user focuses and avoids their phone, the reward is water drops that users can use to tend to a virtual garden, growing and collecting different kinds of plants.

3.1.1 Native tools

Since 2018, both Android and iOS has a feature in the settings of phones to display screen time, and supplied tools for managing it [48], [49]. In both Apple's "Screen Time" and Google's "Digital Wellbeing", the user can see how much time they spent on certain apps, different categories of apps, and the total time [37], [39]. Users

can also block apps during certain periods of the day, or after a certain amount of time has been spent on an app. Both also have different modes, like work mode or bedtime mode, that allows the user to for example block certain apps or show the screen in gray scale during different times of the day. Both Huawei and Samsung also have their own similar versions built in to their smartphones [50], [51]

3.2 Tools within research

Numerous tools for smartphone use have been developed within this field of research, some designed to collect usage data, others to encourage user reflection. These studies use various different categorizations of how participants interact with their phones, some of which were brought up in section 2.4. In this section, we will review several relevant studies and summarize their key findings and results.

3.2.1 Focus on reflection

Lukoff, Yu, Kientz, *et al.* [18] applied a Uses and Gratifications perspective in exploring meaningfulness in smartphone use. With this perspective, they suggest that there is a need for a more user-centered typology for smartphone use. Categories provided by app stores are usually either too vague, or fails to take into account the multiple ways an app can be used. Instead, Lukoff, Yu, Kientz, *et al.* [18] suggest surveys and experiments with users as a better foundation for these kinds of typologies, and propose their own types of smartphone use. Their types were; productivity, information, communication, entertainment, social media (only passive use), and not sure. Using the experience sampling method through a mobile app, they asked users for the motivation behind their smartphone use before, after or during app sessions. The study found that the motivation behind the use and the type it is classified as, can predict how meaningful the user will think it is. Habitual use lead to users feeling a lack of autonomy, and was overall less meaningful.

Terzimehi, Haliburton, Greiner, *et al.* [21] explored the impact of making users reflect on their real world activities while they were using their phone with the app MindPhone. When unlocking their phone, users were faced with a question, either asking about their intention with their phone in that moment, or a question asking which activity the users wanted to do after they had used their phone. The latter proved effective, making the users be reminded of the physical world and the tasks they had to do. They also investigated how writing down thoughts can impact reflections, by making one control group have to type in their answers to the question (active), while the other group simply had to reflect in their head (passive). Both modes had drawbacks, with the active mode being too annoying when users wanted to look up something quick, and the passive mode being too easy to swipe away.

Another recent example of similar ideas being implemented into a physical artifact, is Attention Receipts [52], which is a physical printer that prints a receipt of users' watch session on YouTube, and the total time they spent. Sathya and Nakagaki [52] mention not focusing on limiting the time, but rather increasing reflection about the

quantity and quality of what the user watched. They point out that it is necessary to not view digital well-being as "less screen-time is better" but to rather have a more holistic approach. Furthermore, through their research they found that the materiality of the medium helped to distance the user from the screen and it created more meaning in the data they printed out leading to more meaningful reflection. Additionally, they found that having a digital tool when the consumption also is digital posed some challenges as it could contribute to digital overload. It would add more digital information on top of all the existing digital information that users are already bombarded with [52].

3.2.2 Feature-level evaluation

Cho, Choi, Kim, *et al.* [35] consider intra-app use and also analyzes which features within an app is being used. By developing an analytical system, Finesse, they were able to track feature-level use for four popular apps, and prompt users to evaluate their sessions by selecting which features they regret using. By evaluating which features the users regretted using, the features that felt less meaningful could be pinpointed. The different user scenarios that often lead to less meaningful app use sessions could also be identified. Habitual checking was found to often have little to no reward, which could lead user to get disoriented, lingering on the feature, mindlessly seeking for some rewards [35]. This could also lead to doing a feature-tour, where the user upon not finding enough satisfaction goes through the other features in the app in search for rewards. Another finding was that active form of use, like uploading posts, direct messaging and searching where less regretful than passive forms of use, like viewing stories and posts [35].

Users often felt regretful when they deviated from the original intention they had when they opened the app. However, these deviations were often unconscious or hard to resist as the app's workflow design mixed active and passive features in between each other [35].

Possible rewards found from social media use were social rewards - allowing user to stay connected with friends, informational rewards - the acquisition of new information, personal interests - satisfaction gained from content that corresponds to ones interests, and entertainment rewards - bringing pleasure through fun and entertaining content [35].

Another work that builds on the evaluation of regret on feature-level granularity is FinerMe [36]. Orzikulova, Cho, Chung, *et al.* [36] researched how effective self-reflection contra interventions were and how it differed depending on whether it was on feature-level or on app-level. They developed two versions of their FinerMe app, one that visualized time spent on apps solely on the app level and offered restrictive interventions that restricted the whole app and another that visualized the time spent on different features within an app and offered restrictive interventions for each feature in the app. To limit the scope they chose to only implement it for YouTube and Instagram.

On another note, they chose to not completely block the user from the app or

feature for the rest of the day, as previous studies have found that strong restrictive interventions, despite being effective, caused a lot of frustration and annoyance. Instead, they turned to interaction design friction by creating a 'friction window' which blocked out the user from the app during a one-minute period once the user exceeded their daily limit [36].

Further, Orzikulova, Cho, Chung, *et al.* [36] found that self-reflecting mechanisms did not lead to users reducing their screen time as they were easily bypassable. Instead, they found that restrictive measures on app level significantly reduced the user usage time. However, app-level restriction did not change the amount of time they spent on passive features compared to active ones. In contrast, they found that feature-level restriction lead to users using the passive features less but there was not a significant reduction in screen time spent overall.

3.2.3 Personal visualizations

While not a finished tool, Sosa-Tzec [10] instead explores new ways that screen time can be visualized through 9 sketches of innovative designs. Using analog sketching as a method for exploring delight in design, he presents ideas such as a Tamagotchi-like pet that the user feeds with their mostly used apps, or a lock screen animation with a wildfire that burns more as the user uses their phone. More conventional approaches are also explored, with the most used apps displayed as different sized bubbles, boxes, or stripes. While Sosa-Tzec [10] explores different approaches that represent screen time in a new way, the focus still lies on the total amount of screen time.

Huang, Tory, Aseniero, *et al.* [11] created the research field of personal visualization, visualizations applied in a personal context. This field comes with its own challenges, for example accommodating users with little experience with visualizations, as well as making it fit into their daily life [11]. Schneider, Schauer, Stachl, *et al.* [12] explore how peoples' perception of different kinds of personal visualizations depends on their personality traits. Based on common visualizations of liquid intake on the market, they compared three visualizations; a donut chart, a "living" creature, and a glass of water. The participants first got to use and react to the different visualizations in a lab environment, and then do a personality test. The results of the study was that participants with the personality traits of openness and extroversion to a greater degree preferred the living creature visualization, while participants with the personality trait conscientious, instead had a tendency to deter this visualization.

In a similar study, Wang, Tanahashi, Leaf, *et al.* [13] explores three different versions of visualizations for Facebook data, but does not explore any correlations to personality traits. These visualizations are a lot more rich than the ones from [12], since the data is more complicated, and the visualizations all show multiple parameters at once, for example posts, direct messages and likes. The three visualizations were; a timeline visualization with a stacked area chart, a bar chart and a line chart, then a spark visualization with firework-inspired flares bursting out from the center, and lastly a bouquet of flowers, where the flowers and their petals represent different parameters. The findings of the study were that more abstract visualizations, like the

bouquet and spark visualization in this case, could intrigue the users to explore the data more than the more traditional, in this case the timeline, could. However, they also found that having too much emphasis on aesthetics could lead to users having a harder time connecting the visualization to the data, making it more confusing.

3.3 The research gap

While numerous studies have explored the subject of screen time, there are limitations in prior work. For example, while Cho, Choi, Kim, *et al.* [35] and Orzikulova, Cho, Chung, *et al.* [36] let users categorize their own usage, they focused solely on whether the usage was regretful or not, an approach that lacks nuance. Lukoff, Yu, Kientz, *et al.* [18] did develop a typology for smartphone use with multiple categories, but the focus of the study was not on enabling users to view or reflect on this data themselves. Similarly, Finesse [35] did not visualize or showcase this data back to the user in any way, and while FinerMe [36] did include a bar chart visualization showing daily use of YouTube and Instagram, it did not show how much of that time that the user classified as regretful.

To the best of our knowledge, there have not been any attempts to visualize smartphone use that the user themselves have categorized, and then display it back to the user beyond conventional graphs, such as bar charts. The closest related work is Attention Receipts [52], which display both video titles and time spent on a receipt. While there exists research on more abstract visualizations within the field of personal visualizations, we are not aware of any studies that apply this to visualizations of smartphone use.

4

Methodology

This chapter presents commonly used methodologies, detailing the research approach, different data gathering methods and methods of design. The aim is to provide a rigid framework for the research done. It begins with explaining wicked problems, why they are relevant and how to approach them. In the next section, research through design is discussed and lastly the design process and different applicable methods are presented.

4.1 Wicked problems and design thinking

Smartphone use is a highly complex concept, influencing individuals in ways that are not yet fully understood. The diversity of phone usage and the vast range of content consumed further contribute to this complexity. As an integral part of daily life, smartphone use impacts various aspects, including work, social relationships, and sleep. Given this multifaceted nature of screen time and overuse, the challenges associated with them can be classified as "wicked problems".

Wicked problems are problems that are difficult to define or do not have any clear solutions due to many interdependent factors [53]. Solving wicked problems often requires a deep knowledge of the subject, problem, and all involved stakeholders and requires unique innovative approaches.

Buchanan [54] later connected design thinking to wicked problems and argued how design thinking can help solve wicked problems through its iterative, human-centered approach that emphasizes collaboration, empathy and experimentation. For example, the design process often begins with gathering a deep understanding of the users and stakeholders needs, desires and pain points and through this designers can identify the root causes of these problems and frame them more accurately. This is especially useful for wicked problems, as they can be difficult to define and frame.

Another aspect of design thinking that works well for wicked problems is the iterative cycles of prototyping and testing. The rapid prototyping and testing allows designers to explore multiple solutions and test them in real-world situations, which helps to navigate the uncertainty and adapt to the changing conditions of wicked problems [54]. It is also very common to try to reframe problems by looking at them from different perspectives and redefining them. This allows designers to uncover new opportunities and solutions that they might not have thought of previously. Lastly,

design thinking is very adaptable as the process is very open to change as new insights and discoveries emerge. This is why it enhances the likelihood of devising meaningful and sustainable solutions to complex wicked problems [54].

4.2 Research through design

When using design thinking to address wicked problems, a methodology used to systematically explore, test and validate insights is "research through design". The position of design as research has long been discussed. Zimmerman, Stolterman, and Forlizzi [55] use Frayling's [56] distinctions of 'research about design', 'research for design', and 'research through design'. Research *about* design is research done about the design process itself, and trying to understand the design practice. Research *for* design instead, is done to improve the practice of design, by for example developing frameworks and recommendations. A common practice here is to borrow knowledge and frameworks from other disciplines and applying it to design. Lastly, research *through* design is described as using design as a form of research to explore what the future might look like. Zimmerman, Stolterman, and Forlizzi [55] further build upon this concept of research through design, describing research through design as an approach that can develop theory for design, and theory on design. However, Zimmerman, Stolterman, and Forlizzi [55] also call for a more rigorous process of performing research through design, and a need for more guidelines and protocols. Gaver [57] acknowledges this, but also emphasizes how overly restricting research through design could take away its strongest aspects of challenging the current way of thinking. Instead of trying to make research through design into something it is not, Gaver [57] argues that the design artifacts themselves should be at the core of the work.

4.3 The design process

The double diamond design process takes the well-known design process of divergent and convergent thinking and defines it into a model with four phases, discover, define, develop and deliver [58], see Figure 4.1. For each phase, a set of suitable methods for that specific phase are described.

4.3.1 Discover

In the discover phase, the designer should fully immerse themselves in the topic, and get familiar with the future users' goals, needs, and motivations [58]. It is all about openness and empathy, not being afraid of finding new inspiration, angles and perspectives. In this phase designers can use several different design tools to truly understand the subject and the end user. They can use design methods such as field studies, observations, interviews, competitive testing and literature reviews [60].

Interviews can be good to collect firsthand experiences, perceptions, opinions and attitudes from end users. There are four main types of interviews: open-ended or unstructured, structured, semi-structured and group interviews [61, Chapter 8]. The

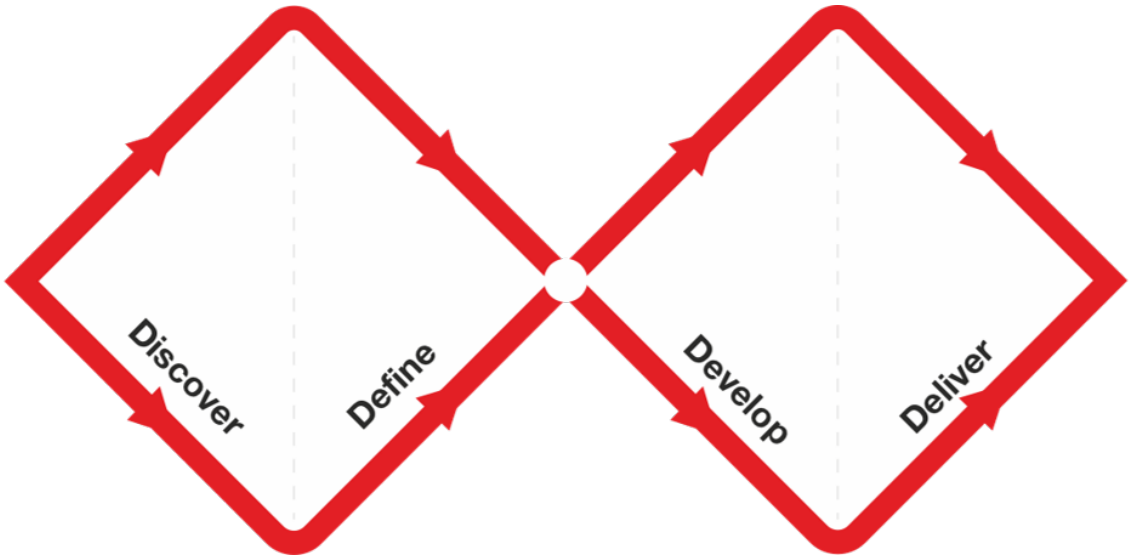


Figure 4.1: The double diamond design process. Source: [59]

format that is the most appropriate depends on what the purpose of the interviews are. In the discover phase, unstructured or semi-structured interviews might be preferred as they give room for being more exploratory and allows flexible detours. Additionally, they are often more conversational and comfortable for participants [60]. In contrast, structured interviews is better for more specific goals, for example if the goal is to receive feedback on a specific design feature [61, Chapter 8].

An alternative or complement to interviews can be questionnaires. While being simple to produce, it is still important to pay attention to the amount of questions and how they are worded [60]. Questions can be more open-ended to receive more qualitative data, or more close-ended to get more quantitative data that is easier to analyze [60].

Another fundamental research skill is observation, which requires attentive looking and systematic recording of phenomena [60]. It can be of people, artifacts, environment, events, behaviors and interactions. Observations can also be either semi-structured, casual or more structured and systematic [60]. Similarly to interviews, more casual and semi-structured observations are used in the more exploratory phase of the design process where the goal is to gain information through immersion, especially in territories that are new to the designer. More structured or systematic observations utilizing worksheets and checklists or other forms of coding are better where environmental or behavioral elements are targeted and well defined.

Competitive testing and literature reviews can also be useful to gain valuable insights and information on what works well and what works poorly [60]. By the end of the discover phase the designer should have a thorough understanding of the context of the challenge and the potential end user.

4.3.2 Define

In the second phase, the focus lies on analyzing the data and insights that has been gathered during the discover phase. It is all about connecting the dots, finding common themes, spotting patterns and searching for insights in order to be able to clearly frame the design challenge [58]. Here, tools that help to organize information and finding common patterns in user behaviors such as Thematic networks or Affinity diagrams are useful [60].

Thematic networks helps to identify, organize and connect the most common themes in rich qualitative data. It helps the designer to synthesize the data they have gathered into meaningful, actionable design insights. By splitting up the themes into basic themes, clustering the basic themes into organizing themes and then finding the overarching global themes the most unifying message can be identified [60].

Affinity diagramming is a technique used to externalize and meaningfully cluster observations and insights from research [60]. This method helps designers organize their findings and stay grounded in data throughout the design process. By writing key findings, observations, and requirements on sticky notes, analyzing them individually, and then grouping them based on similarities, designers can identify common themes derived from the data they have gathered [60].

4.3.3 Develop

During the develop phase, the diagram diverges again, and the goal with the phase is to invent many new solutions, iterate and test.

Brainstorming is a way of generating many ideas in a short time, and while it is common to use the word casually, Kelley and Littman [62, Chapter 4] explains brainstorming as a skill that one can improve. He also gives some suggestions on how to improve a brainstorming session, for example by numbering your ideas, getting physical and visual, and having playful rules, like not critiquing or debating ideas. One method for brainstorming is Crazy 8, a method where each participants sketch eight ideas in eight minutes, with the goal being quantity over quality [63].

Connected to brainstorming, an integral tool to the design process is drawing. Drawing can be used during many parts of the process and for a multitude of purposes, like examining and understanding the world, exploring and generating ideas, and explaining ideas and concepts to other people [60]. Sosa-Tzec [10] brings up sketching specifically in relation to digital well-being, explaining the value of using mediums that makes the designer be patient and embrace imperfections.

Personas are profiles, created based on information about real users [60]. While not being examples of actual users, they are representative of a specific user group with certain traits and experiences. A persona has a name, picture, description of their life and goals, and relevant behaviors connected to the design in question. They are a way for the designer to design for specific users, rather than trying to design for everyone.

Scenarios are a method for envisioning the future use of what is being designed, by writing a narrative of a person using the product or service [60]. This helps the designer stay grounded in the users' needs and avoid only designing towards the technical requirements. Scenarios are often written from the perspective of personas.

Similarly to scenarios, storyboards use narratives to help designers understand potential ways users interact with the artifact being designed [60]. Storyboards use graphics and drawings to visualize different scenarios and in which context the product can be used. Truong, Hayes, and Abowd [64] suggest guidelines for designing storyboards, that include; use text, but be aware that the words have influence, include people if their reaction is important, indicate passage of time only if it is important, and generally use minimal detail, as long as the point still is clear.

The MoSCoW method is a prioritization technique where all features of a future product is listed, and then categorized into the following categories, where it also gets its name from; Must have, Should have, Could have, and Won't have (this time). This is a way to figure out what is absolutely needed for the product to exist, and what is less important and can be cut away if the time is limited [65].

Prototypes are a partial representation of the design, and can be both high fidelity, very close to the finished product, or low fidelity, in the early stages of the process. [66, Chapter 10]. Lim, Stolterman, and Tenenberg [67] view prototypes as filters, proposing that the incomplete part of the prototype is just as helpful as the complete part. Prototypes as filters means that different aspects of the design idea can be shown through different prototypes, and it is up to the the designer to design prototypes that show the qualities of interest. Lim, Stolterman, and Tenenberg [67] also discuss a mixed-fidelity approach to prototyping, prototypes that are low-fidelity in some aspects and high-fidelity in others, depending on what the aim is with the prototype. Generally though, low fidelity prototypes are built early on in the design process, quickly and with cheap materials [66]. Paper prototypes are a common and popular way to quickly try out ideas without using much time or resources. High fidelity prototypes on the other hand, look and feel very similar to the finished product, and are developed quite late in the process. Benyon, Turner, and Turner [66] mention that it is very important that the high fidelity prototype is believable, without any errors.

4.3.4 Deliver

During the last stage in the double diamond design process, a decision has been made on pursuing one solution, and this solution is further iterated on and tested [58].

Usability testing is done to evaluate how usable a product is, and is commonly done in a laboratory, with an observation room where researchers can observe the participants during the test [61, Chapter 15]. However, this does not allow designers to see users in their own environment, which can be important in order to test some products. Field studies can instead be a way to observe users in their natural settings, and the data is recorded through observation and interviews, through video, photos,

audio, and notes [61, Chapter 15]. A method inspired by experience method sampling can also be used, where participants can be interrupted throughout their day and be asked to fill out a diary, for example on their smartphones [61, Chapter 15]. The advantage of using smartphones for this type of study is for instance the ability to use the built-in sensors in order to collect more data on the context surrounding the entry [68].

Questionnaires and interviews can also be used here, with the purpose of evaluating and getting feedback on what has been created.

If there are multiple versions of an interface that should be tested, the between-subjects or within-subjects approaches can be chosen [69]. For between-subjects, each participant only tests one of the versions, while for within-subjects, each participant tests both versions, switching in the middle of the test. There are pros and cons with both approaches, where between-subjects can be shorter and easier to set up, and also minimizes the risk of the participants learning the first version, which can make them more efficient in the second one. On the other hand, within-subjects requires less participants, and will also bring their experiences to both versions, which minimizes noise.

5

Timeplan

In this chapter, a timeplan outlining each phase of the project is presented. Based on the methods described in the previous chapter, this chapter discusses how they can be applied to this particular project.

5.1 Prestudy

The time plan is divided into several steps. Based on the double diamond work process, the first four weeks were planned to be spent in the exploratory discover phase, with the focus on exploring the market and research area as much as possible. This phase, referred to as the pre-study, would include a literature review, testing of screen time apps on the market and conducting interviews. The interviews would be semi-structured, focusing on letting the participants reflect on their own screen time data, as well as their thoughts on the apps they use and how they use them. Participants were to be recruited through student forums, and the interviews would be recorded. Afterwards, the recordings would be transcribed and then analyzed together with the other findings from the pre-study through the thematic networks method. The outcomes from the discover phase were then to be distilled into a planning report, where the scope and aim of the project would be decided in the define stage.

5.2 First iteration

The development of the app was planned to be divided into two iterations, with the first iteration focusing on brainstorming and exploring several different ideas in the develop phase. Here, sketching, scenarios and storyboards would be useful to generate many ideas. During the deliver phase, one or a few low fidelity prototypes would be made, that would be tested and evaluated on a small scale before the second iteration is begun.

5.3 Second iteration

Based on the lessons learned from the first iteration, the second iteration would repeat the develop and deliver stage. The second iteration was also planned to

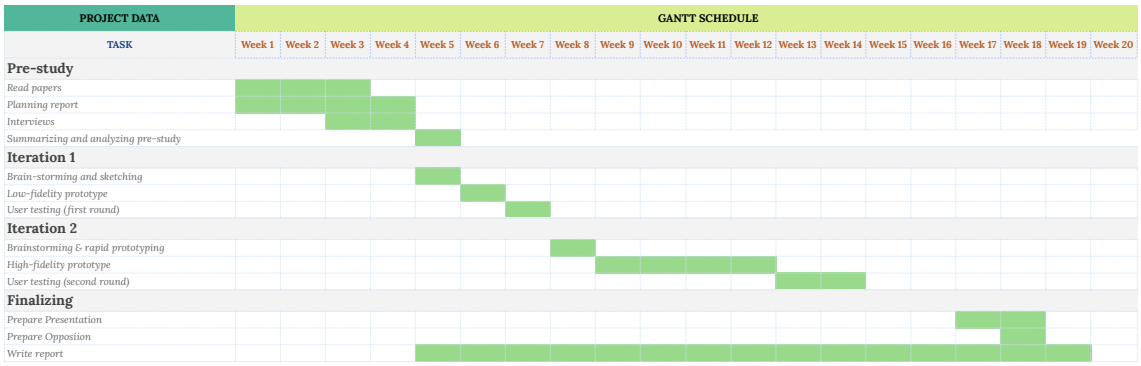


Figure 5.1: Gantt Schema of the time plan

be longer, as a more high fidelity prototype would be created, that should work for conducting more extensive tests. The high fidelity prototype would then be tested more rigorously with participants using the prototype for about a week, and data could be gathered through a similar way to the experience sampling method. Depending on the outcome of the iterations, multiple variants of the prototype could be tested, with a few days allocated to each variant. After the testing period, interviews would be held, focusing on the experience of using the prototype and possible changes in behavior and reflection.

After the prototype and testing would be done, a few weeks were planned to be dedicated only to writing and finalizing the report and preparing for the final presentation. It was also planned to write on the report and to keep a working diary throughout the entire project, documenting decisions and the process. During the entire project weekly or bi-weekly meetings would be held with our supervisor to ensure everything goes well and stays on track.

6

Execution

This chapter will go through the entire design and implementation process, from pre-study and low-fidelity prototyping to evaluating the finished prototype.

6.1 Pre-study

The pre-study was carried out during the first five weeks of the thesis. It included doing a literature review, reading relevant previous research and theories, going through and testing existing tools on the market as well as performing interviews to understand the user group better and to gain valuable design insights that helped to lay the foundation for the coming design work.

6.1.1 Interviews inquiring about views on screen time

It was decided to use interviews instead of questionnaires in the pre-study, since the goal was to get more insight into the user group and their reflections about their screen time, and it would be easier to get this type of information through interviews. As said in section 4.3.1, semi-structured interviews can be a good option for the exploratory phase of the project, since it allows for both some consistency between interviews, while still being open enough that there is possibility to dig deeper and ask follow-up questions during the interviews. With this approach, interview questions were formulated and discussed with the supervisor. The final questions used can be found in the appendix, A.1. To gather interview participants, a spreadsheet with available times was sent out to forums for software engineering students at Chalmers University of Technology. Seven participants filled in this form, and interviews were held. After a meeting with our supervisor, we decided on gathering a few more participants, preferably also with more diverse occupations, not only students at our own division. We did not feel we had the time to send out another form with possible times, so we utilized our own connections of friends and family, to get more participants with other occupations. Through this convenience sampling approach we managed to have seven more interviews, leading to a total of 14 interviews. A list of all interviewees can be seen in table 6.1. Every interview except one were held in Swedish, and the remaining one was held in English. All interviews were recorded, and later transcribed. In the beginning, the interviews were transcribed manually, but some interviews were also transcribed with the AI tool whisper [70], with executables from [71]. This tool can be run locally, which

means no data had to be uploaded to any external sites. While this tool made the transcribing process faster, we still listened through the interviews, checking and improving the generated transcription afterwards.

Interviewee	Age	Occupation
$I1$	23	student
$I2$	24	student
$I3$	22	student
$I4$	21	student
$I5$	28	student
$I6$	24	student
$I7$	31	student
$I8$	24	student
$I9$	27	student
$I10$	25	working
$I11$	23	working
$I12$	21	student
$I13$	29	working
$I14$	23	student

Table 6.1: The participating interviewees of the prestudy.

6.1.2 Thematic analysis

After the interviews had been held, the focus was to analyze the data and gather valuable insights from it. This was done through the method Thematic networks, see 4.3.2. We started by finding basic themes, or codes, from the transcribed interviews. This was done in a free, open-source program called Taguette [72] that we downloaded and hosted on our own local server. We coded half of the interviews each and then went over the other person's coded interviews to see that the codes set were reasonable and added more if the other person had missed anything.

After that was done, all the codes were written down on post-it notes. When coding the interviews we also divided some of them into smaller sub-codes. These were written down on smaller post-it notes. All the notes were then grouped and organized into larger organizing themes on a whiteboard. see fig 6.1. These were then further grouped and organized into larger global themes in a separate document.

6.1.3 Resulting themes from the pre-study

The pre-study resulted in 5 global themes based on 14 organizing themes, that are presented below, together with excerpts from the interviews. Most of the excerpts have been translated from Swedish. The global themes are; *Designed for engagement*, *Compare to others*, *Hard to resist*, *Conscious of screen time*, and *Enriches life*. The corresponding organizing themes for each global theme can be viewed in the appendix, see A.3.

Figure 6.1: Post-it notes with codes grouped into organizing themes

Designed for engagement

In the interviews, participants expressed how they often got stuck scrolling for longer than they wanted to, because of how apps were designed. With many apps having an endless feed of content, the desire to feel finished with an app is sometimes impossible. *I*13 explains:

> "*It's really made so that you get completely hooked on it. And I don't like the ads, and I don't like many of the suggestions that come up. It would feel better if it were just the people you follow, and then the feed ends.*."

As many apps have multiple features for different purposes, some participants felt they still needed or wanted an app because they used some features, but they overused some other features, wishing they could take them away:

> "*But I don't really know how to get around it. Like, there have been times when I've uninstalled YouTube, but then it's like, now I can't watch YouTube which actually also has long form content. I would've liked to remove YouTube Shorts, I hate YouTube Shorts.*. (*I*2)"

Compare to others

When participants used social media they often felt discontent with their own lives as they easily compared themselves to others. Social media users can portray their life as being better than it actually is, and others that compare themselves with it

can end up feeling bad that their life is not as perfect:

> "*This computer, the one you can turn into an iPad... If I had it, I would've passed all my exams. I mean, I've passed my exams without it, so clearly I can pass exams without it. ... Or this car. Im creating a dream life for myself. One I don't actually need, but that I'm jealous of.* (*I*14)"

A focus on gaining likes and followers can also lead to feeling bad about oneself, as *I*8 describes:

> "*Instagram, it's...I don't like how it can turn into comparing yourself to other peoples posts. 'Ah, you got like 120 likes and I only got 100.' It can get pretty toxic really fast, I feel.* "

Hard to resist

With smartphones being a necessary tool to get through everyday life, and something most people always has with them, they can be hard to resist. Participants expressed regret for using their phone when they wished they did something else, embracing a moment in nature, being more present with their partner, or when they had something important to do.

> "*I always get distracted. It doesn't have to be that I actively check it. But, for example, if I'm having coffee with someone and it's lying there, and then I get a notification 'Oh, it's that person, I have to deal with this', and by then I've already lost the thread between me and the person I'm having coffee with. So when I meet people, I think it's unnecessary to bring my phone.* (*I*14)"

While it can be a distracting, participants also expressed that their smartphone can be convenient to use when they are bored or want a break, as it is easily accessible:

> "*I almost always am on my phone when commuting, just because it's a good distraction, and the alternative is just kind of staring out the window and thinking.* (*I*7)"

Conscious of screen time

Users can be more or less conscious of how much time they spend on their phone. Most of the interviewed people had a pretty good understanding of what they spent their time on and were not surprised about their data. There were however some confusion and frustration around how it was visualized in the built in visualization on their phone:

> "*Sometimes I read books on my phone before going to bed, instead of using the Kindle. And then it's like, now your screen time has spiked by an extra hour and a half. I don't know, maybe it's that screen time has really negative connotations, like it's supposed to be a bad thing. But in a case like that, if I'm using it to read a book, it's very different from me*

> *watching TikTok for two hours. Then it's actually a positive thing that I've been reading. And yet I still have high screen time.* ($I2$)"

Some felt unsatisfied with their use and tried to actively restrict and reduce it using different strategies and tools, while others felt satisfied and did not try to reduce it. There were also participants that felt unsatisfied but did not try to do anything about it. Examples of strategies that participants used to restrict their phone was; deleting apps, setting app timers and reminders, and increasing friction, for example by making apps harder to access. However, these were not always effective, as $I1$ describes:

> "*I've realized that it helps me, or well, it makes me feel bad about the time I spend on it, like if I'm having a bad day and sitting at home on my phone for a long time, I see the minutes ticking up and feel a bit guilty. But then it doesn't really help me stop using it.*"

Enriches life

Smartphones are not only stealing time, participants felt that they were necessary and useful in their daily life:

> "*That I put things in the calendar. Yeah, planning my day, setting reminders. I also use it as an alarm clock. I take a lot of notes on it. I also feel like it's really good if you want answers to questions.* ($I10$)"

Additionally, many participants used their smartphone for things that made their lives better. Being connected with others and communicating was something that many participants regarded as meaningful. Furthermore, using apps to get inspiration from others, learning new things and engaging in hobbies are all reasons that smartphones are still such a big part of peoples lives. $I9$ describes this as:

> "*I have some friends I've known for a really long time who sometimes post new pictures or stories, and I've started to enjoy that, you want to know that your old friends are still doing well, that things are going well for them... now someone has kids, someone's traveling all the time, someone's just enjoying reading books, it's just fun sometimes to know.*"

6.2 First Iteration

6.2.1 Personas and scenarios based on the interviews

Based on the data from the interviews, three personas were made, seen in fig 6.2. These personas are a representation of multiple users, with different reoccurring characteristics from the interviews. The first persona, Linda, feels a lot of regret over using her phone too much, especially in social settings, and finds it hard to uninstall apps since she still likes some features of them. She also fears she will miss out on things if she deletes them. She likes engaging in her hobbies through social media, but feels she sometimes compares herself to the people she follows. Having tried app timers, she often feels like they do not help her.

The second persona, Thomas, likes to keep track of his screen time, since he used to get stuck scrolling before, and has since uninstalled apps in order to be more mindful of his time. Although, because he has uninstalled a lot of his social media apps, he often feels disconnected from society and he feels like he misses out on things happening in his friends' lives. Sometimes he does fall back into old habits of scrolling but overall feels very satisfied with his screen time. He mostly uses his phones for finding new recipes and other things he's interested in.

The third persona, Rose, usually uses her phone for reading. She likes to hike and when she is home or outdoors she actively puts her phone away in order to stay in the moment. She also feels like screen time often have such a negative view from society but she herself does not feel like her screen time is bad or ill spent. Usually, it is mainly from her reading.

Creating the personas helped to further concretize the different findings from the interviews and helped to tie the themes together. It gave us concrete users to base our further design work on. After the personas were made we also created different scenarios, both positive and negative experiences each persona could have when using their phones. This helped to identify different use cases and situations that the designs could be grounded in. These can be found in the appendix, see A.2.

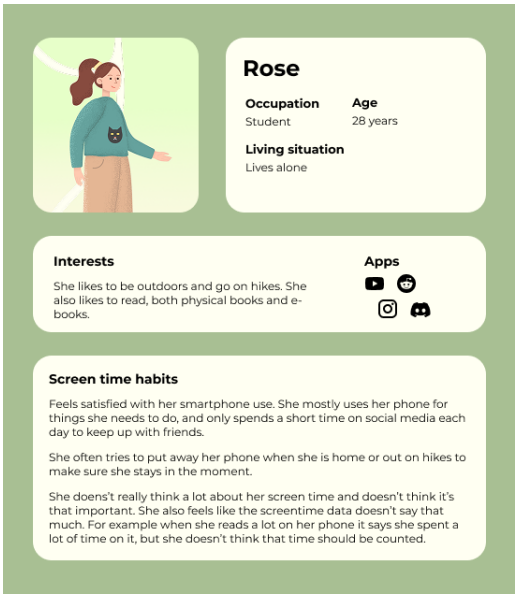
6.2.2 Brainstorming of possible concepts

With a good understanding of the potential users, we started brainstorming, mainly focusing on generating many ideas. We used the method Crazy 8 [63] and once we had a few ideas we continued developing those, sketching storyboards and wireframes. Some of the sketches generated during this process can be seen in figure 6.3. This helped us find problems with the ideas and we could discuss potential solutions. After a few days of exploring ideas, we had two main ideas. One tool with a focus on visualization, where a session of smartphone use is represented by a flower field, with a different type of flower for each category of use. The users would categorize their use themselves, through a prompt after an app had been used for a certain amount of time, or after exiting an app. These categories would be based on the themes found during our pre-study, with also an option to classify the session as meaningless, or not worth their time. This would result in a dying flower, or something similar. This visualization of a meadow would be showcased on the lock screen of the phone, so the users can remember to look at it. While inside the app, each flower could be pressed to see what category of use it represents, the length of the session, and which app was used. Other kinds of visualizations were also considered, for example, visualizing each app session as different kinds of buildings in a small building complex or city, as different fruits in a fruit tree, or as different fishes in a fish tank.

The other app idea was a more gamified tool, where the smartphone use would be represented by a plant companion. The user would rate their use after each session, and the plant would feel bad and start to whither if the phone was used in a less meaningful way from the users perspective, otherwise the plant would continue growing. Different plants would take varying amount of days to grow, and once finished, the user could unlock or buy a new plant to grow. The plants could have

(a) Persona 1: Linda

(b) Persona 2: Thomas



(c) Persona 3: Rose

Figure 6.2: Personas

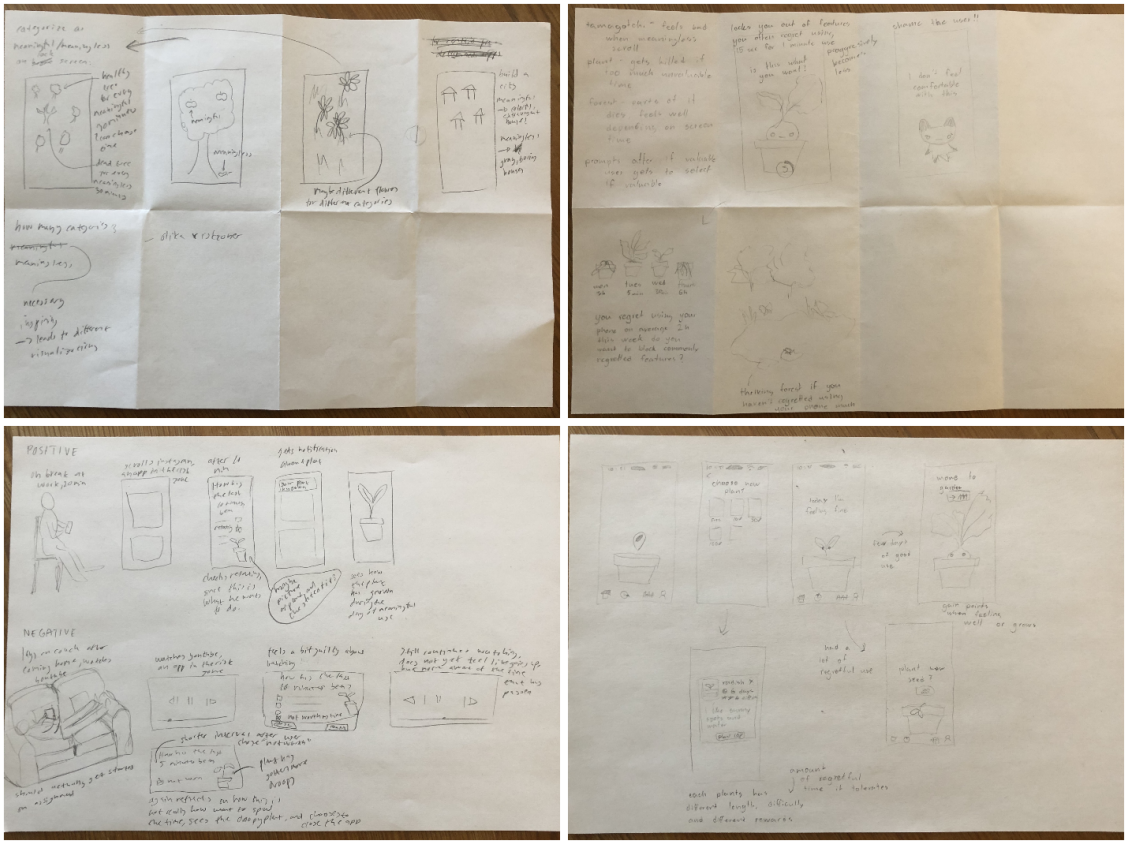


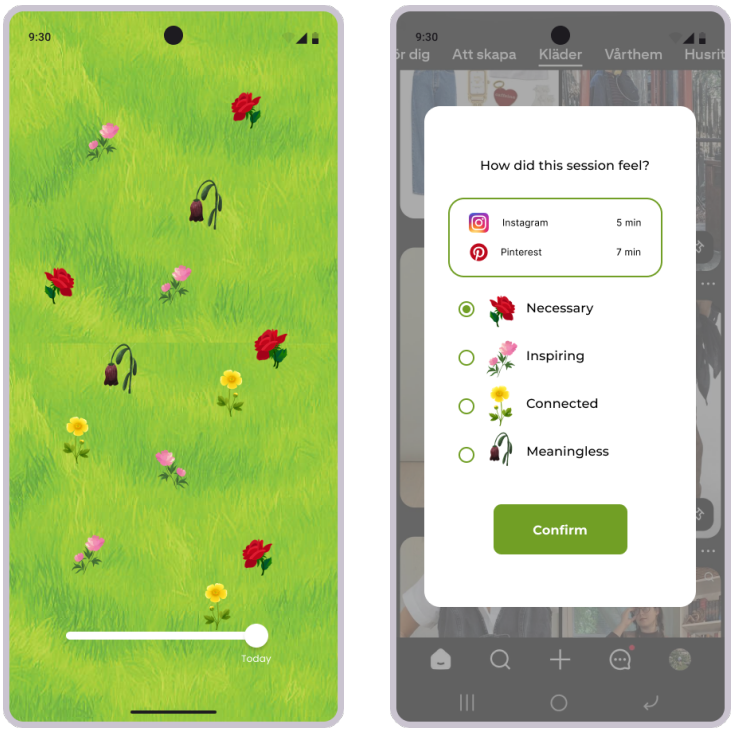
Figure 6.3: Sketches from brainstorming session

difficulty levels, with some being more resistant to "bad" smartphone use, so that users themselves could decide how strict they want to be.

We were interested in this metaphor of a living thing, specifically plants, representing smartphone use, partly because it is a metaphor that we have seen before in this context, such as Flora [46], Focus Plant [47] and Forest [38]. Another reason is because the connotations that comes with plants as something to take care of, and care about, that can grow and improve, or whither and die.

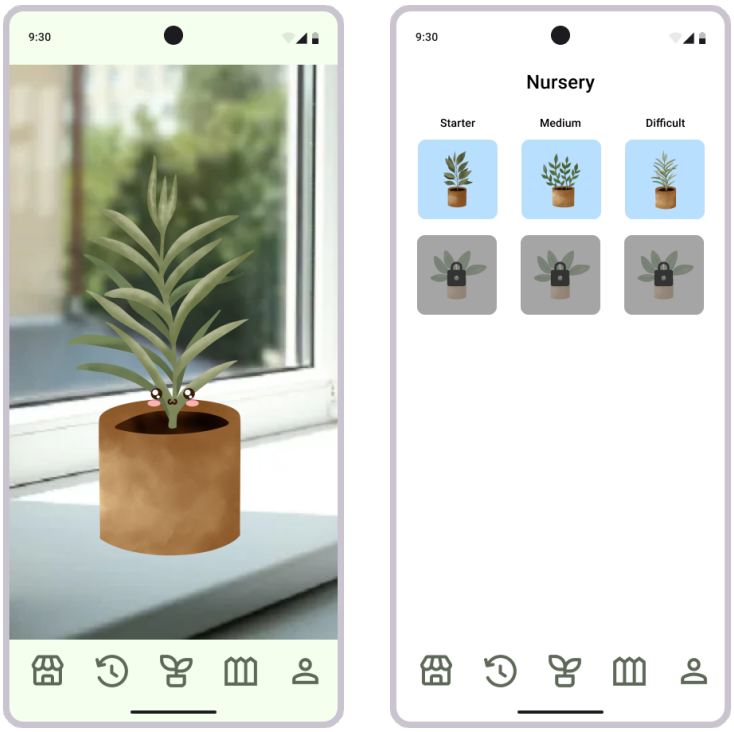
6.2.3 Low fidelity prototypes of two ideas

To explore both ideas further, we had a lot of discussions, continued sketching and figuring out how both apps would work. We tried to envision how our different personas would use them and what would benefit or bother them. Furthermore, we developed mock-ups of the main functionalities in Figma, see figure 6.4 and figure 6.5. We showed both ideas to some peers and to our supervisor, and after further considerations, we decided to work with the flower application. The reason for this was that it felt more novel and new, and further connected to the themes from the pre-study. Many participants in the interviews expressed that many apps could be used for multiple purposes, and spark different feelings. This would be better reflected in the flower application, with its different flowers for different feelings, rather than the more black-and-white approach of the plant application.



(a) Home page (b) Prompt

Figure 6.4: Mock-up screens for Flower app.



(a) Home page (b) Get new plant

Figure 6.5: Mock-up screens for Plant app.

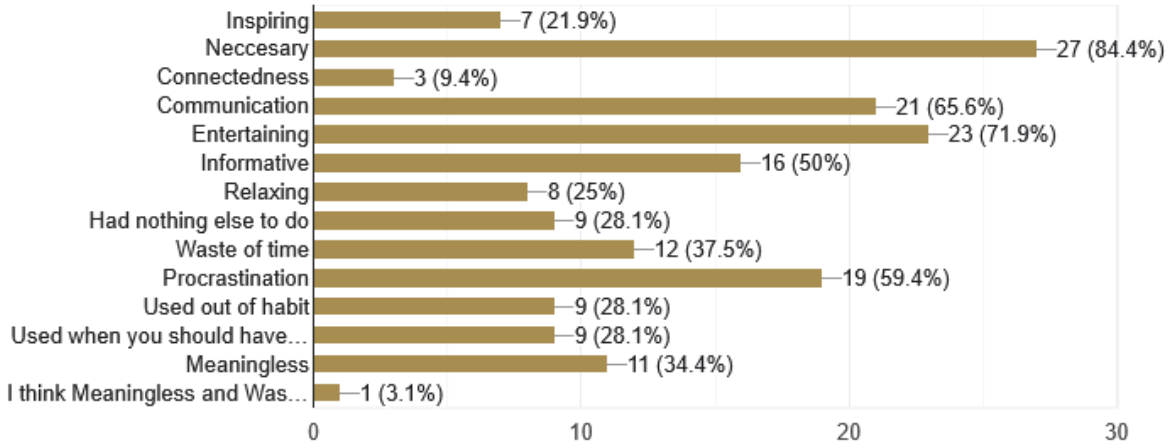


Figure 6.6: Preferred categories to include, from the questionnaire

6.2.4 Questionnaire for feedback on concept

After the concept to focus on had been chosen, we further developed the flower field visualization mock-up, defining more specifics of how it would work. Additionally a form was sent out in our student division forum to gather feedback and general opinions on the idea, how automated people would prefer it to be, what categories they found suitable, how to best visualize more negative type of screen time and any other thoughts they might have. When deciding which categories to include in the survey we decided to base them on the themes found in the prestudy, for example some people said they used their phone to pass time, for relaxation, for finding information, to get inspired or for staying connected, etc.

The survey received 32 responses in total and it was clear that most people preferred to have the app be quite automated. Automated in this context being the possibility to set default categories for some apps, and be prompted to choose a category after each use for other apps. On a scale from 1 to 5, 1 being fully automated and 5 being fully manual, 19 chose a 2. Only one person chose 1, and one person chose 5, indicating most people wanted a middle ground. One person suggested that the app should learn what categories the user usually categorized an app as at different times of the day and based on that, automatically choose a category for the session without prompting the user. This could definitely be a future improvement, but was a bit out of scope for this project. The results of which categories were most relevant can be seen in figure 6.6, with the top categories being necessary, entertaining, communication and procrastination. Lastly, participants' preferred visualizations for less meaningful smartphone use can be seen in figure 6.7, where the most popular visualization was dying flowers. Some also suggested that they should be able to choose their own categories or choose which flower represented which category. This was something we took into account later when the app was developed.

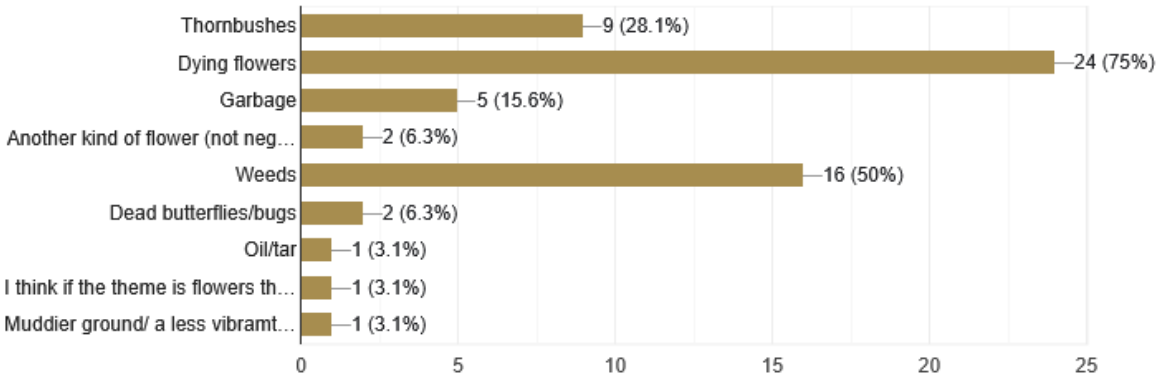


Figure 6.7: Preferred visual representations of less meaningful smartphone use, from the questionnaire

6.3 Second iteration

Once we had gotten feedback on the idea, the development process began. While our original idea was just to implement the flower field app, after a while we came to the conclusion that there were two quite distinct parts of the idea that we wanted to explore. First, the self-categorization aspect, where users themselves put a category on the app session they had, but also the visualization aspect where we wanted to represent their data in a new way, with a flower field. This lead us to develop two versions of the app, with the same categorization logic in the background, but having one more conventional visualization in the form of a typical bar graph alongside the flower field.

6.3.1 MoSCoW prioritization technique

In order to figure out in what order the different functionality should be developed, and to figure out the scope of the app, we did the MoSCoW prioritization technique, see section 4.3.3. Features such as, the flowers representing the time spent in each app, and being able to view previous days were essential, or must haves, while features such as having different themes or more customization were less important for this version of the project. This also helped us to talk about which features we both saw as important, and to take into account features that had been suggested by participants in the survey.

6.3.2 Development tools

The app itself was developed with Kotlin, in Android Studio. The reason for this was that Kotlin is more suited for Android development in comparison to Java [73]. Another possible choice for smartphone app development, that is also cross-platform, could have been React Native or Flutter, but since we needed access to more device-specific functionality, we opted to develop only for Android. For the UI, we used Jetpack Compose, which is the recommended approach from Android

[74]. It has built in access to Material Design components, which makes it easier to have consistent UI. We also used Material design and their Material Theme Builder to handle color themes. For debugging the app we used both an emulator and a physical android device. For version control, we used GitHub [75].

6.3.3 Database setup

To store the data that was gathered we decided that using a local SQL database would be the best option. This has the advantage that everything is stored locally on the phone and the app does not need to rely on internet connection. It also makes it easier for the user to feel like they are in control of their data. To build on this we also decided that no data should get sent to us automatically. Instead, it is the user themselves that send it to us at the end of the test through a share function in the app.

The database itself was setup using Room [76], a library in Android that helps you work with a SQLite database, but in a easier and safer way. With regards to the comments we got on the survey, we chose to store the categories themselves, and which flower they correlated to in the database, to make it easy to implement for users to add, edit or remove categories themselves in the future. It also makes it easy to update data when needed, and helps keep related information connected and saved properly.

Also stored in the database is a table of all the apps that are downloaded on the phone, their name, package name and icon. This is needed in order to be able to display which app was used and to be able to set default categories on them. There is another table for the default categories containing the app's package name and the default category for that app. Lastly, there is the table which stores all the app usage sessions, it contains the package name of the used app, the start time of the session, the duration, the category it was categorized as, and the date the session happened. Instead of trying to group several sessions together or try to create longer sessions with several apps and store them that way in the database, we chose to store the sessions one by one, this is because it is easier to tweak or change how it could be grouped or visualized later.

6.3.4 The tracking service

In order to track when an app is used, Android's accessibility service [77] is utilized. This feature gives access to tracking accessibility events, for example when the state of the window is changed, which can then be used to track which app the user currently is viewing. Actually intended for tools for accessibility, such as screen readers, it has been used by other screen time tools, such as Finesse [35] and FinerMe [36] to track features and app usage. There are two downsides with this approach. Since the feature could be used for malicious purposes, the user has to manually turn on permission for it in settings. For this same reason, more secure apps such as BankID and some bank apps, do not allow this service to be running in the background, and it has to be temporarily turned off while using those apps.

In order to try to ensure that the actual tracking corresponded to the way the user used their phone, some precautions were taken. For example, subtracting the time the phone was locked from the session, and also only tracking launchable apps, so not any background services were tracked. One example of such an app was the keyboard app, which made it look like the keyboard app was used instead of the app the keyboard was used in.

6.3.5 Development of the overlay prompt

Each time the user has exited an app, an overlay prompt should pop up, to prompt the user to categorize the session they just had. It turned to be some compatibility issues between the accessibility service and the library we used for the UI, Jetpack Compose. The compose view needed a life cycle to handle when the view was added and disposed, but this was not provided by the accessibility service. To solve this, we created our own life cycle class and attached it to the view. While this might not be the most conventional solution, it made it possible for us to use Jetpack Compose for the overlay as well, so that the UI could be consistent throughout the entire app.

6.3.6 Development of the flower field feature

The flower field was set up in Jetpack Compose using a Canvas component in order to be able to add filters and overlays. This enabled it to mimic different light settings on the field and made it look more realistic and cohesive. It also added the possibility to add different lightning depending on the time of the day. This was something we did not have time to implement before the tests, but it is something that could be potentially added in the future.

The illustrations for both the background and the flowers themselves were entirely made by us using Procreate [78]. We had a lot of discussions on how to visualize the different sessions as flowers. For example, should one flower visualize a "whole" session of smartphone usage with several apps, or should one flower represent one app session? How would we differentiate shorter sessions from longer sessions, should flowers have different sizes depending on length of time, or should longer sessions be visualized with a small cluster of several flowers? In the end, we chose for each flower to represent one app session and that the size would correlate to the length of the session as we thought it best represented how many sessions it actually was. If you had clusters of flowers for longer sessions this could be interpreted as many small sessions. For this we decided to include three different sizes of flowers representing three different time intervals. The different time intervals were 10 seconds to 1 minute for small flowers, 1 minute to 10 minutes for medium flowers and then 10 minutes or longer for large flowers. We chose not to include sessions shorter than 10 seconds as there would be too many small sessions that might not matter very much, and to make sure the user does not get prompted too frequently. We chose to set the longest interval to 10 minutes or longer, as a single app is not frequently used for a long time uninterrupted, even if the phone is used for an extensive amount of time. For example, the average TikTok session is around 10 minutes long while other social media sessions are usually shorter, around 1 to 5 minutes long [79].

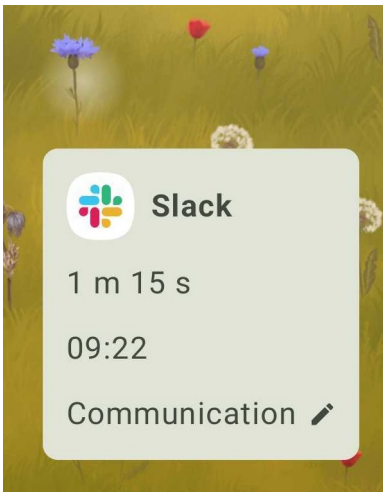


Figure 6.8: The info box on the flower field visualization

We also wanted to be able to include information in the flower field about each session, for example which app was used, when it was used and how long it was used. This was achieved by making the flowers tappable, showing a small info box when you tap them, see fig 6.8. We considered whether the smaller flowers should be tappable at all or if they would be too small or insignificant, but in the end we chose to make them tappable as well.

6.3.7 Development of the bar chart

The bar chart visualization was created in order to act as a more conventional version of the app, where bar charts are a common way to display screen time, and that many users recognize from the built-in screen time visualization in their phones. The bar chart showed the total time for each category that has been used that day. Compared to the flower field visualization, the bar chart was a lot easier to implement, as it is something that is more standardized and many people has done before.

6.3.8 The chosen categories

Likewise to Lukoff, Yu, Kientz, *et al.* [18] we chose to base our categories from the findings in our pre-study. The chosen categories were based on the themes found in the interviews and the answers from the form that was sent out. The seven most popular categories from the form were included; Necessary, Entertaining, Informative, Communication, Passing time, Procrastination, Waste of time and one that appeared a lot in the interviews, Inspiring. When it came to choosing which types of flower to illustrate, we tried to find flower types that naturally grow in meadows and fields, and had distinct colors and shapes to make sure they were easily distinguishable even for people with colorblindness. We also tried to match the flowers to the different categories, which all can be seen in figure 6.2. For example, the necessary category was chosen to be represented by a neutral white cow parsley flower, while the more positive categories; informative, communication, entertaining

and inspiring were more colorful. The more neutral passing time category were represented by a dandelion clock, and the more negative categories, procrastination and waste of time, were depicted as withering flowers.

In the end, there were eight categories, with eight correlating types of flowers. Each flower type also had three different sizes with two variants of the middle size to add variation, so in total 32 different flowers were illustrated.

Description	**s (10s-1m)**	**m (1m-10m)**	**l (>10m)**
Necessary			
Entertaining			
Communication			
Informative			
Inspiring			
Passing time			
Procrastination			
Waste of time			

Table 6.2: The corresponding flowers to the different categories.

6.3.9 Development of the lock screen wallpaper

In order to remind the user to look at the data often, it was important to be able to set the flower field as the lock screen wallpaper and update it whenever a new flower was added. This turned out to be easier said than done as there are certain

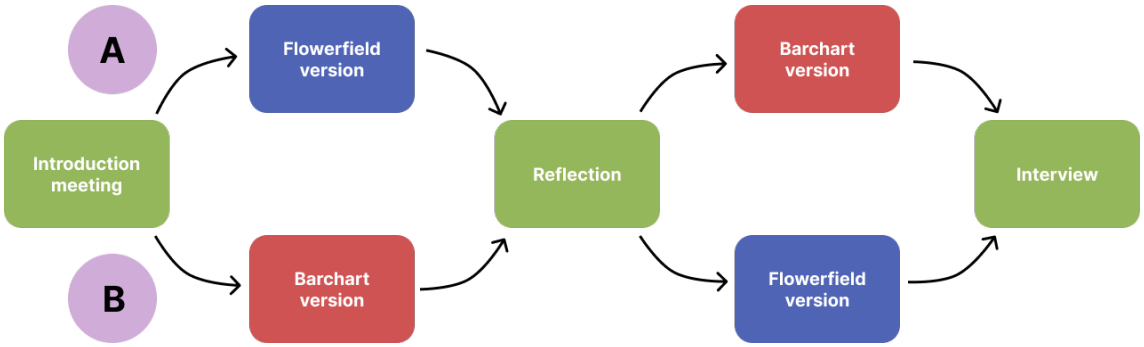


Figure 6.9: User testing flow.

limitations when setting the wallpaper. For example, special permissions are needed, and on most devices it can only be set from a foreground service. Additionally, in order to set the wallpaper there needs to be an image to set it to, so the flower field compose first had to be rendered into an image bitmap, which also needs to be done in the foreground in order for it to render correctly. This was solved by creating a foreground service that was only active long enough to render and set the lock screen wallpaper and was completely invisible to the user. Whenever there was a change in the app sessions table in the database, the foreground service was started from the accessibility service. Furthermore, in order to only set the lock screen wallpaper if the user wanted to, we included it as a setting in the settings page, which was turned off as the default.

6.3.10 User testing and interviews

After the second iteration, we did our main round of user testing. Since there were two versions of the app, the within-subjects approach [69] was used. This approach was chosen instead of the between-subjects, largely because it was difficult to find participants, and within-subjects requires fewer people. Participants were gathered from those who participated in the pre-study, those who volunteered in the questionnaire, and by convenience sampling. For all participants we had an introductory meeting, where we introduced the project, what was expected of them, how we handled their data, and how the test would work. We intentionally did not tell the participants too much about the flower visualization or what the different types of flower and sizes represented. The reason for this was that we wanted to see how intuitive the visualization was. We then guided them through installing the app, and made sure it worked properly on their phones. Half of the participants began using the bar chart version of the app, and the other half the flower field version. After half of the test period had passed (4 days), we sent them instructions on how to switch version of the app, and also asked them to fill in a form with their reflections from the first period. After the entire 8 days had passed, we booked in an interview in the following week after they had finished using the app. These interviews were semi-structured and the questions can be seen in the appendix, see A.4. All interviews except for two were held in Swedish. See figure 6.9 for an overview of the process.

After the interviews, the data was processed in a similar way as in the pre-study. The interviews were transcribed using the AI tool whisper, which can be ran locally on our own computers. We then listened through the recordings again, and made needed adjustments to the generated transcriptions. The interviews and the reflections were coded in Taguette, and then grouped into clusters digitally with FigJam [80].

6.4 Time plan evaluation

In this section, the differences between the intended time plan and the actual outcomes will be discussed. Firstly, the interviews in the pre-study were planned to take two weeks, but since we got a lower response to the initial message we sent out, it took more time to find more participants. The interview portion also became more extensive than what initially was planned, which took a longer time, but also yielded better results. This lead to the brainstorming phase starting a week later, but we were still able to begin working on the high fidelity prototype at the time that was planned. The testing was started half a week earlier than planned, mainly because we realized we had allocated too little time for this. With our testing period being eight days, two weeks would not leave enough time to both book introductory meetings with participants, hold interviews, transcribe interviews, and analyze the data. Even though the testing phase was started early, it still took longer than planned. The biggest reason for this was that the app did not work for two participants, the reason being that they used an older version of Android than we had tested on. While we did spend a few days on debugging this error, we were not able to solve it, and unexpectedly had to find a few more participants. Although some phases of the project were underestimated and a few unforeseen events occurred, other phases had been overestimated, leaving us with margins that allowed us to complete everything we had originally planned.

7

Results

In this chapter the results from the study is presented. First, there is an overview of the resulting app 'Flower' that was used during the user testing and the bugs and problems it had. Then the results of the user tests, the gathered data and insights from the halftime reflection and interviews will be presented. This includes an overview of how participants categorized their apps, the effects of the self-categorization process, their general impressions of the visualizations, and their preferences, including which visualization they favored and the reasons behind their choices.

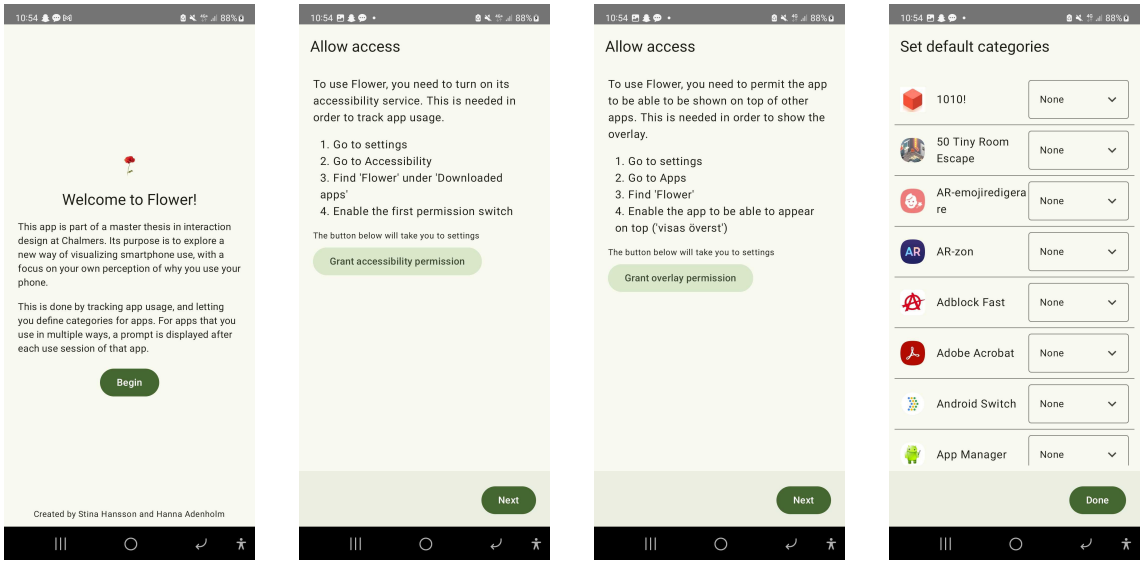
7.1 The Flower application

7.1.1 Setup

The first time the app is opened, the user is faced with an introduction page, with information about the project and the app, see figure 7.1a. Next, they are prompted to allow access to the accessibility service that tracks apps, as well as permission for the overlay to be shown above other apps, see figure 7.1b and 7.1c. Third, they can fill in which default categories they want to have, figure 7.1d. While we during our introduction session guided users through this process, this gave them some extra information and support to setup the app properly.

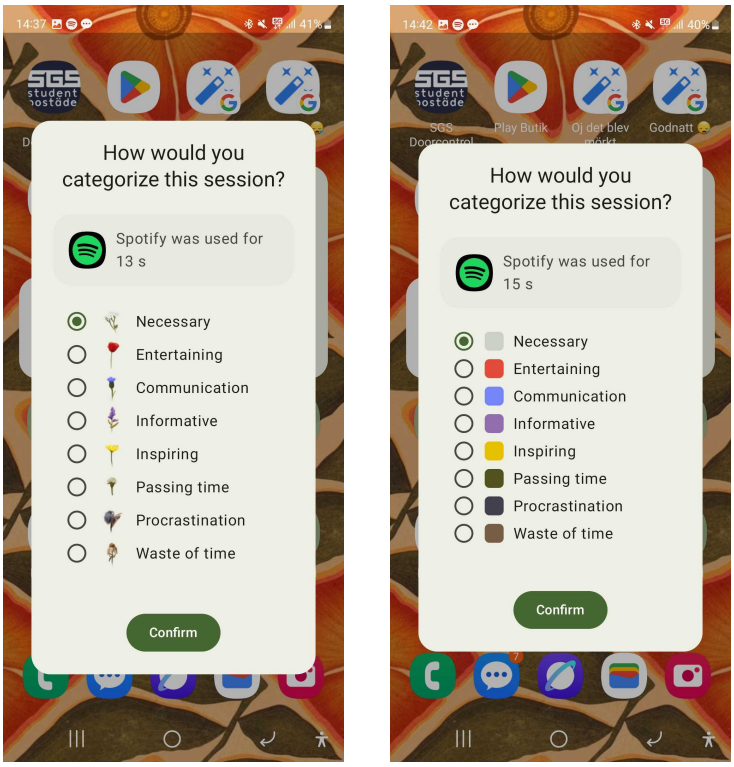
7.1.2 The overlay prompt

For apps with no default category, when exiting the app, the overlay prompt shows up, asking the user to categorize the session they just had. If the user clicks outside the box, it is closed, but the session is still logged, as the default option "Necessary". While the category might not be correct, it ensures that no session is accidentally deleted if the user misclicks. The overlay prompt could be designed so that it is less easy to accidentally dismiss, but as the overlay shows up over everything else on the phone, we felt it was more important that the user did not feel stuck, and that it could be easily closed. The flower field version of the overlay can be seen in figure 7.2a, while the bar chart version, that has colors instead of flowers for the categories, can be seen in figure 7.2b.



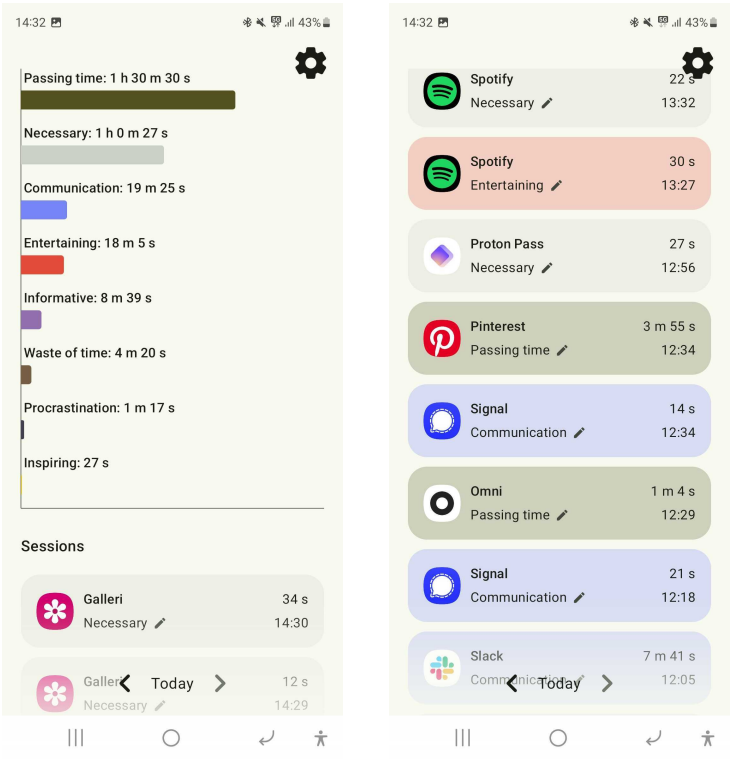
(a) Welcome (b) Accessibility service access (c) Overlay access (d) Default categories

Figure 7.1: Setup



(a) Flower field version (b) Bar chart version

Figure 7.2: Overlay



(a) Home (b) Sessions

Figure 7.3: Bar chart

7.1.3 The bar chart

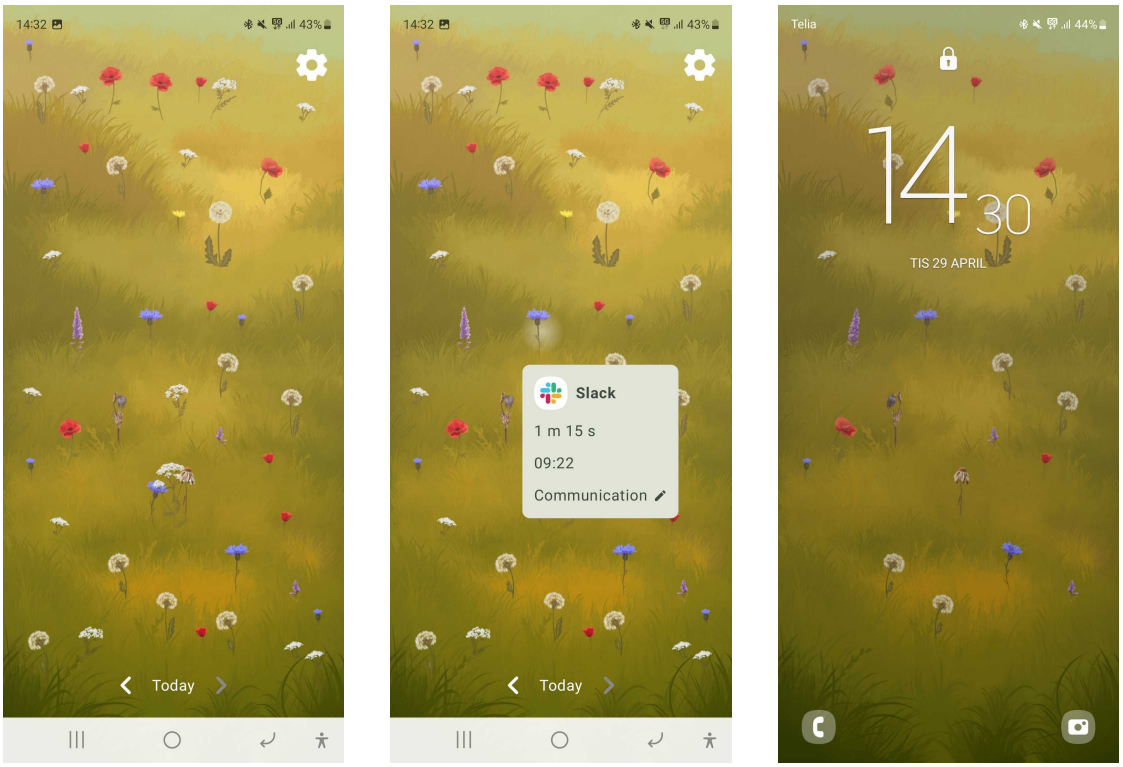
The bar chart view shows the total time in each category, see figure 7.3a, and has a list of all the sessions below the graph, see figure 7.3b. These session cards are color coded according to the categories, and the colors for the bars are similar in color to the corresponding flower in the flower field view, to make it easier for the user when they switch version of the app.

7.1.4 The flower field

The flower field view represents all app sessions as flowers, see figure 7.4a, where the type of flower is different depending on which category it has, and the size corresponds to the length of the session. Every flower can be tapped to receive more information about the session, such as the duration, the start time and the app that was used, see figure 7.4b. The flower field can also be set as the lock screen for the phone, see figure 7.4c, and is updated every time there is a new session.

7.1.5 Settings

The settings page mainly holds the default categories for the apps, see figure 7.5b, and the possibility to switch between the two versions of the app, see figure 7.5a. There is also the option for setting the flower field as the lock screen wallpaper, and a way to send the app sessions as an csv, which is the easiest way for us the



(a) Home (b) Infobox (c) Lock screen

Figure 7.4: Flower field

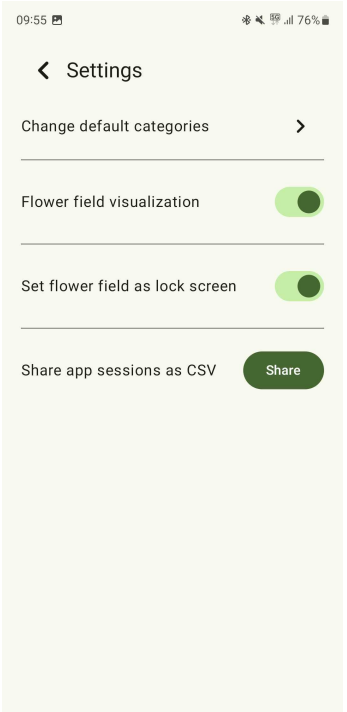
receive the data from the participants' phones. When setting default categories, the user can choose between all the existing categories, or "None", in which the user is prompted via the overlay for that app. If a category is chosen for an app, those app sessions will be automatically logged as that category. There is also an option "Ignore this app", which is for apps that may run in the background and give faulty sessions, and the user can then choose to avoid logging those apps.

7.1.6 Editing and deleting

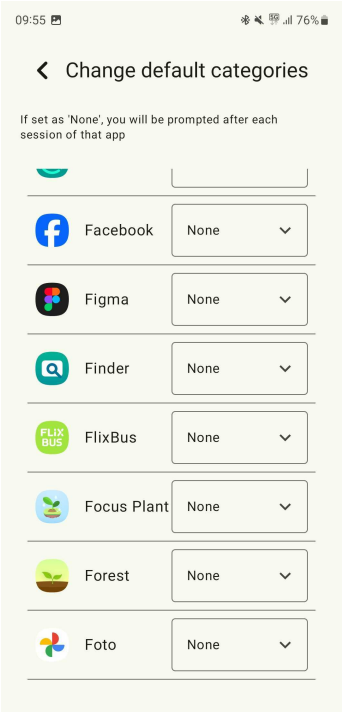
As a consequence of the overlay prompt being quite easy to dismiss, there is a risk of the user accidentally saving a session as the wrong category. For this reason, we implemented a function to be able to change the category of a previous session, see figure 7.6. Similarly, if the app malfunctions and accidentally logs the session as being too long, there is also the option of deleting the session entirely. While it would be good if these functions were not necessary, it does allow the users to make their data more representative, if something goes wrong, or if they change their mind after choosing a category.

7.1.7 Bugs and problems

One thing that became apparent during the tests was that the app, and especially the tracking service, worked differently on different android phones. While the logging

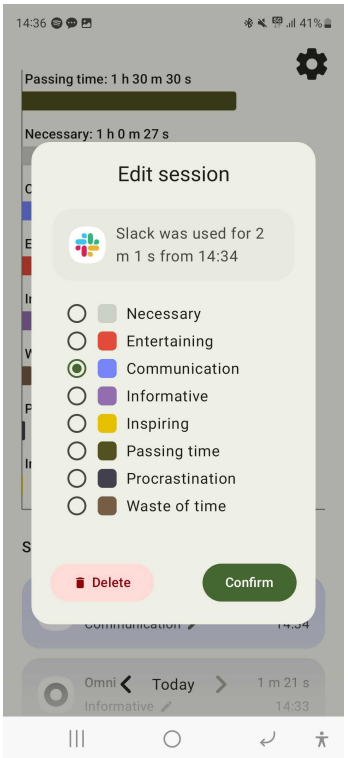


(a) Settings page

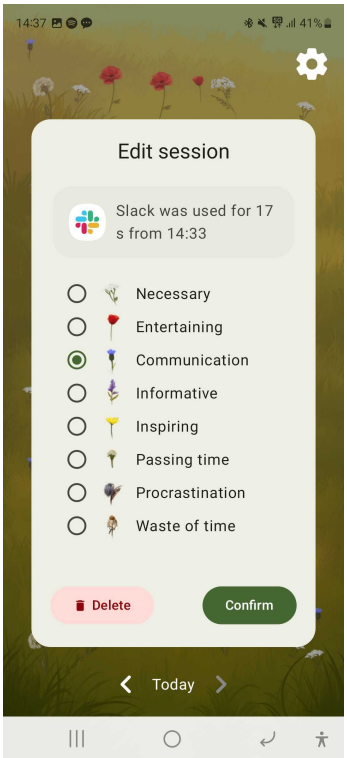


(b) Default categories

Figure 7.5: Settings



(a) Bar chart version



(b) Flower field version

Figure 7.6: Edit and delete

worked as intended on the devices used during the development phase, with the exception of one or two apps that ran in the background and gave incorrect sessions lengths, it turned out to be more problems for some participants. Many of the participants had troubles with the tracking of sessions sometimes continuing even when the phone was locked, which led to sessions that were longer than they had used the app for. While the edit and delete functions existed for users to adjust faulty sessions, users could not adjust the time, which led to data that did not represent the actual usage. Some participants had more problems than others, ranging from it working as intended, to them getting the wrong session length every time they locked their phone while being inside an app. This lead to both wrong data, but also to unpredicted behavior from users changing the way they used their phone so that the tracking service could track the sessions more accurately. For example, one participant mentioned that they made sure to close apps properly before locking their phone, in order to avoid faulty sessions.

Some participants also reported that the overlay occasionally popped up even when they had not used the app. For apps that were rarely used, the "ignore this app" option could remedy this problem. Another problem was that for some apps, the session was interrupted while the app was still in use, leading to one long session instead being represented by many small ones. The reason for this is because of the way the accessibility service is used, by listening to specific events and then detecting which app is in use. Some actions, like rotating the screen or taking a screenshot can lead to there being a new event, and the tracking service thinks the session has ended.

One person also had problems with the app crashing every time a default category was set. We managed to get it working so that the rest of the app worked as intended, and the person could still perform the test, but had to choose a category after each session they used their phone.

Unfortunately, these bugs means that the data is less trustworthy, especially with sessions being generally longer than they were in reality. For this reason, we will mostly focus on the users' categorization of sessions, rather than the sessions' frequency and length.

7.2 Results from user testing

In this section the results from the user testing is presented. Data gathered from the user testing is shown, together with themes and qualitative data from the interviews and midpoint reflections. Participants are anonymized, and will be referred to as P1-P12, with either a letter A or B depending on which version of the app they began the test with. Group A used the flower field version first, while group B used the bar chart version first. In table 7.1 the participants numbers can be seen, if they participated in the interviews, their age and what bugs or problems they had. A majority of the quotes below have been translated from Swedish.

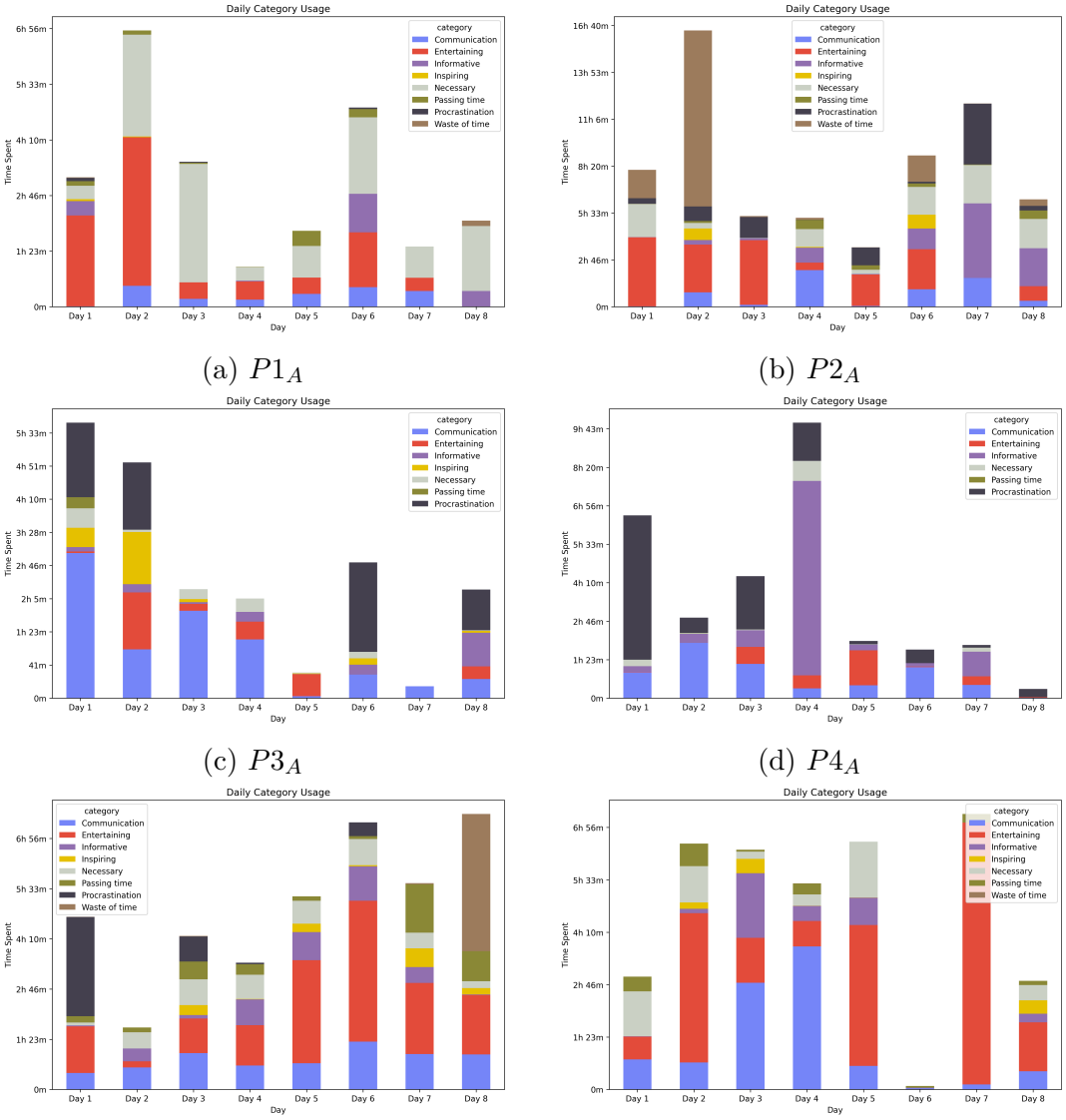
Participant	Age	Participated in pre-study	Bugs & problems
$P1_A$	23	Yes, as $I1$	Sometimes logged when locked, tried removing faulty sessions
$P2_A$	19	No	Often logged when locked, did not remove faulty sessions
$P3_A$	20	No	Misunderstood and only used Flower viz, no experienced problems with logging.
$P4_A$	21	Yes, as $I4$	No experienced problems with logging
$P5_A$	28	Yes, as $I5$	Could not set defaults, often logged when locked, tried removing faulty sessions
$P6_A$	24	Yes, as $I6$	Sometimes logged when locked, did not remove faulty sessions
$P7_B$	31	Yes, as $I7$	Sometimes logged when locked, tried removing faulty sessions
$P8_B$	23	No	Sometimes logged when locked, sometimes removed faulty sessions
$P9_B$	27	Yes, as $I9$	Sometimes logged when locked, did not remove faulty sessions
$P10_B$	27	No	Sometimes logged when locked, diligently removed faulty sessions
$P11_B$	22	No	Seldom logged when locked, sometimes did not log after unlocking, tried removing faulty sessions
$P12_B$	23	No	No experienced problems with logging

Table 7.1: Overview of user test participants and their reported issues

7.2.1 Participants' categorization

The categorization part of the app was almost identical no matter the version that was used. For each app, the user had the option of setting a default category, or leaving it blank, which would make the app prompt for a category after each time that app was used. Many appreciated the default category feature, as it allowed for less interruptions while using their phone, and as they would have pressed the same category each time.

The collected data clearly indicates that the amount of time each person has used the different categories varies significantly. This can be seen in figure 7.7 and 7.8 depicting the participants' daily category use. Some participants, for example $P10_B$, used more of the negative categories, procrastination and waste of time, while others barely used them, for example $P8_A$. Many of the participants also reported that they did not use the inspiring category at all, which can also be seen in the graphs. Meanwhile, $P9_B$ used inspiring almost half of their logged time. This showcases how the categories participants used varied from person to person. For example, $P10_B$ and $P11_B$ used passing time much more compared to the other participants, $P8_B$ categorized the majority as communication, and $P1_A$ instead categorized many



(a) $P1_A$ (b) $P2_A$

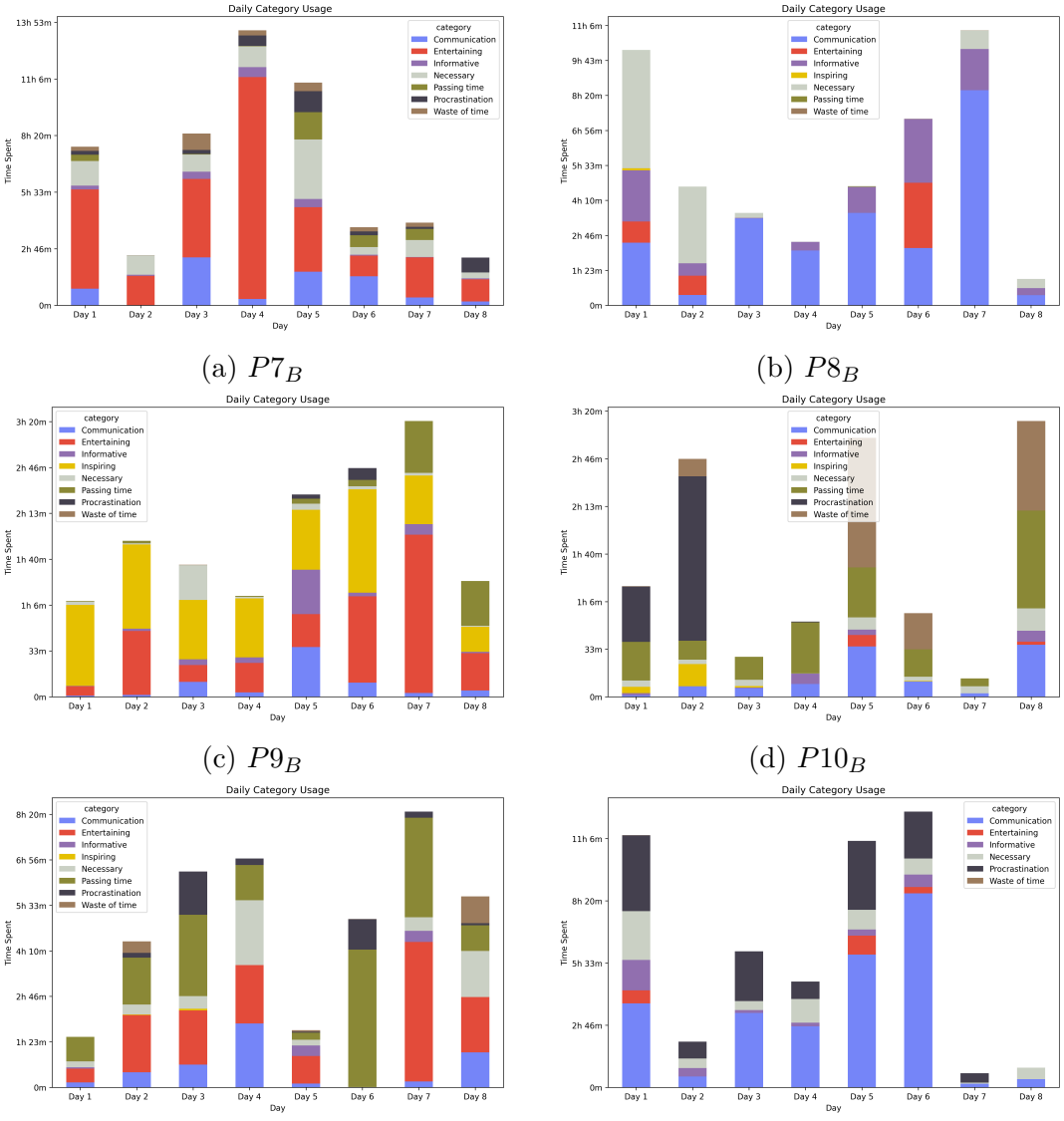
(c) $P3_A$ (d) $P4_A$

(e) $P5_A$ (f) $P6_A$

Figure 7.7: Daily category use, test group A

sessions as necessary.

An important thing to note though is that as there were some bugs that sometimes made the duration of the sessions longer than they were supposed to, the data that was gathered is not entirely accurate. Some participants did not have this problem at all, while others struggled with it more (more information on how correct their data is can be seen in table 7.1). In addition, participants could have forgotten to turn on the accessibility service again after using a more secure app like BankID and having to turn off the service, which may lead to gaps in the data (see section 7.1.7). While this is something that should be kept in mind when analyzing the data, the categorization aspect of the data is more accurate. Another thing to note is that the alarm clock app was removed from the data as several of the participants mentioned it logging time when it should not have.



(a) $P7_B$ (b) $P8_B$

(c) $P9_B$ (d) $P10_B$

(e) $P11_B$ (f) $P12_B$

Figure 7.8: Daily category use, test group B

Besides varying between different participants it is clear in the data that the categorization varies between days, which some participants also witnessed:

> "*It's also fun because you can see how it varies from day to day. Depending on whether you've used it a lot at school, it might mostly be just passing time. Whereas if you're more at home, it might be more about entertainment. Procrastination is present every day in my life. It's always there.*" ($P5_A$)

$P9_B$ reflected about the difference between weekend and workdays and how they used more entertaining and inspiring categories during the weekend and then on the Monday after categorized more sessions as necessary.

Most interviewees felt that choosing a category could be both easy or difficult, depending on the app and the context of using it. Single purposes apps, like many communication apps, was said to be easier to choose category for rather than apps that had multiple purposes:

> "*So for them, I felt that apps like YouTube are more ambiguous, usually the ones that don't have a specific purpose can be more... then they're either maybe entertaining, informative or something, if it felt good, or if it felt bad then it was a passing time or procrastination. The apps that don't have a purpose were more context-dependent in terms of what you categorize them as, whereas some apps, this one you use for bus tickets, this one you use for communication, so it was pretty easy to assign those.*" ($P1_A$)

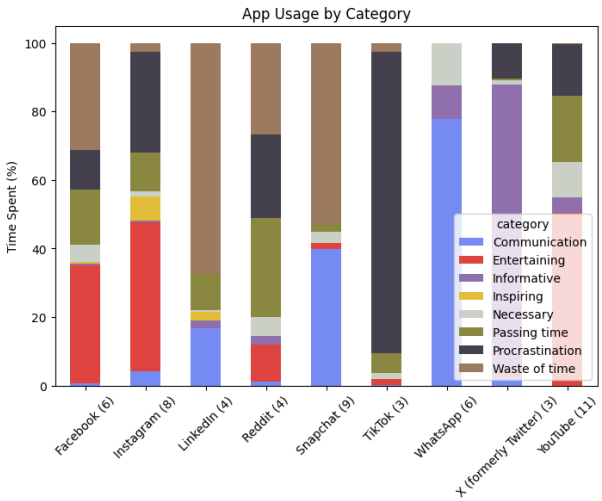
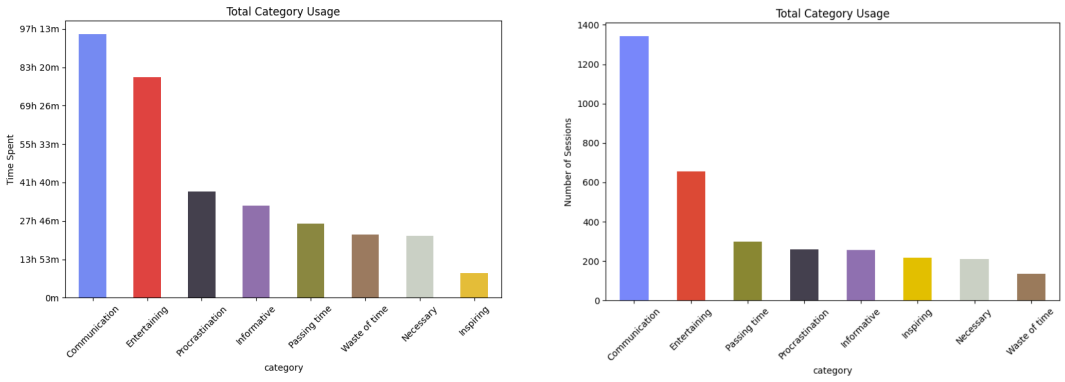


Figure 7.9: Participants' categorization of social media

Social media platforms frequently emerged as a category of apps that served multiple purposes, making them more difficult to classify. Figure 7.9 summarizes how all participants categorized various social media apps used during the study. Beside the app name there is also a number indicating how many of the participants that used the app during the test. It can be seen that some of them varied considerably, for example Facebook, Instagram and Youtube, while others had mostly one single

category, for example TikTok being mostly procrastination, X mostly informative, WhatsApp mostly communication and LinkedIn mostly a waste of time. These categorizations should not be generalized, as the number of participants was limited and no single app was used by all individuals. The way each app was categorized is likely to have been strongly influenced by the individual perspectives of the participants.

Another interesting observation is that the communication category accounted for the highest overall usage. This is illustrated in the two graphs in figure 7.10 that depicts how much each category was used in total, one based on the total amount of time and the other based on total number of sessions. It is clear that communication and entertaining were the most used while the other categories varies a bit depending on if counted by time or number of sessions. The differences between the two graphs also suggest that session lengths vary considerably depending on the category of use. For example, communication has almost double the amount of sessions compared to entertaining, but time-vise they are pretty similar. This indicates that communication sessions tend to be shorter and entertainment sessions tend to be longer. The same goes for procrastination, the procrastination sessions also tend to be longer.



(a) Time (b) Number of sessions

Figure 7.10: Most used categories, first by time, then by number of sessions

While some participants felt the existing categories were sufficient, many wanted more options or the ability to add their own. Suggestions included additional subjective categories, such as "productive" or a category for habitual use. Others suggested categories that referred more to the content or purpose of the app, rather than the feeling it evoked, for example, "transport", "education", "sports", and "food":

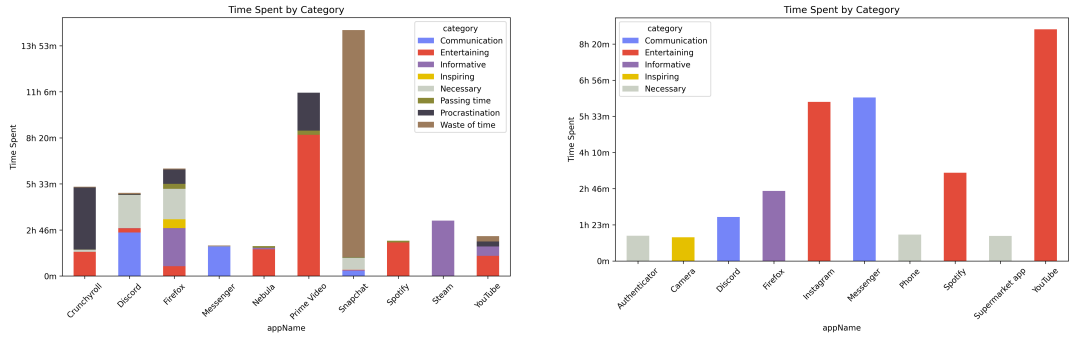
> "*I would have added more categories, ... I would have added a hobby category, for when you're like writing, it's productive but it's not just entertaining, even though it is like a hobby. Or some kind of actual productive category, because sometimes I do some like work on my phone, for like my job, I order invoices, and I'm on my phone to do that, so that is work. I think maybe some kind of like, food category too, if you're just ordering stuff, or looking at menus, that would also be good, cause it's not really entertainment, but it is not browsing either*" ($P7_B$)

A few participants expressed that they could feel that a category was missing, but that they could not pinpoint what they would have categorized it as. $P5_A$ describes:

> "*It's not like I had another category where I thought, this one is missing. So I've still felt that if I don't know why I've spent this time, then I need to think a bit more about what it actually was that I did. And why I can't fit it into any category. So I think the categories themselves make you reflect on what I'm actually doing.*"

7.2.2 The effects of self-categorizing

Something that became apparent during the interviews was that some participants used default categories for most of the apps they used during the test. An example of this can be seen in figure 7.11 that showcases two examples of how the participants top ten most used app were categorized. $P2_A$ purposefully did not set default categories, leading to multiple different categories for most apps, while $P6_A$ set default categories on many apps, which meant that most apps had only been categorized as one category during the entire test. The graphs of the rest of the participants top ten used apps can be found in the appendix A.5 where similar patterns can be seen for participants $P3_A$, $P4_A$, $P6_A$, $P8_B$, $P9_B$ and $P12_B$ who also used many default categories. Three of these participants reported that they did reflect more on their phone use, however, three of them said they did not.



(a) $P2_A$ (b) $P6_A$

Figure 7.11: Examples of categorization of the participants' top 10 most used apps

In contrast, the other six participants, participant $P1_A$, $P2_A$, $P7_B$, $P10_B$ and $P11_B$, said in their interviews that they did not set as many default categories ($P5_A$ also had a bug that did not allow them to set any). This also seems to be reflected in their data as many of the apps they used have been categorized with several different categories, see $P2_A$ in fig 7.11 and the rest in the appendix, see A.5. Participant $P2_A$, $P7_B$, $P10_B$ and $P11_B$ all said that they purposefully did not set many default categories as they wanted to reflect more and wanted the categorization of the app to represent what they use the app for, not what the app was made for. For example, $P7_B$ said

> *"I didn't set too many default, just for things I knew would be entertainment, for example, I didn't put discord as communication, even though*

> *that is what I primarily use it for, because sometimes I would just use it as a time waster. For like chrome and stuff, even though I use it all the time I wanted it specifically to ask me and prompt me, because I didnt want to generalize everything."*

All six of $P1_A$, $P2_A$, $P5_A$, $P7_B$, $P10_B$ and $P11_B$ also said that they thought that categorizing the sessions after they had used them was important for their reflection. They reflected more about why they were using an app in a certain moment, which could depend on what they used the app for, but also if it was appropriate for the context, for example if they were at home or in public. It was often while categorizing sessions that they reflected the most:

> *"Later it became more that I didnt go into the app that much, but mostly thought about it when I was choosing the category. What am I doing right now? For example, if I'm on the tram, I'm passing time on my phone. Whereas at home, maybe I'm entertaining myself."* ($P2_A$)

$P5_A$ also explained an increased reflection about their use even before the overlay popped up:

> "*Now I feel like I'm always thinking about which category to assign to each session, which makes me think about why I'm doing something while I'm doing it.*"

In some cases, seeing the categorization overlay was enough to make some participants put down their phone when they felt they had used it too much, but many also mentioned that their behavior did not change in any way from using the app.

However, some participants also felt that the overlay was annoying, and that it could lead to getting distracted and maybe in the long run feeling a bit fatigued of using the application:

> "*And add, in some way, more distraction to the use of the phone. That you have to think more, you might get a bit exhausted from it or something. So there are pros and cons to using it, I think. There were a lot of things that were interesting, but some things were maybe a bit so-so, kind of awkward or whatever you'd call it.*" ($P1_A$)

7.2.3 The visualizations' effect on the participants

By letting the users themselves categorize their data, the visualizations becomes more representative of how they actually use their phone compared to other screen time apps. $P12_B$ describes it as following:

> "*Because you often get these notifications on your phone like: This was your weekly screen time, you know? And then you react like, wow, three hours. Am I really on my phone for three hours? Oh my god, thats a lot of the day. But then when you look at your app, it's like, okay okay, one and a half hours spent online. And all I do online is search for information. Well, I've been studying, so that makes sense. Maybe*

> *it's not so unreasonable to use the phone for three hours a day. Youre looking things up and so on.*"

Many participants also explained how it was easier to see their negative use, and a few participants were surprised on how much of their phone use they classed as a negative category. On the other hand, a few participants instead saw how they used their phone in a way they felt was good, and could then see this represented in the visualizations.

The majority of participants preferred the flower field visualization, most arguing it was more aesthetic or gave a more positive view on their phone use. In participant group A that started with the flower field, all the participants said they preferred the flower visualization except $P2_A$ that said they would have liked a combination of both visualizations. There was also $P3_A$ that misunderstood the test and only used the flower version so they did not have an opinion on the bar chart version. In test group B, however, the opinions where more split. Half of the participants, $P9_B$, $P10_B$ and $P11_B$, said they preferred the bar chart version, while the other half, $P7_B$, $P8_B$ and $P13_B$, preferred the flower field. The opinions varied in strength as well, some did not like the bar chart version at all and barely used it, while some had more vague opinions, liking both versions but having a small preference.

7.2.3.1 Participants' thoughts on the bar chart

Participants that enjoyed using the bar chart said that they liked the familiarity of it, and that the information was clearly presented:

> "*The bar chart is really good because it has this kind of universal feedback or appearance. Here you have a clear comparison, this bar is bigger than the other one. ... And also the text, it's like reading bar charts in math when you're little, so you just kind of have that with you already.*" ($P9_B$)

Many, including some of those who preferred the flower field, also liked the aspect of being able to see the total time, which is not visible in the same way in the flower field. In a more practical aspect, some participants also said that it was easier to find sessions they wanted to edit or delete in the bar chart version than for the flower field.

This familiarity of the bar chart was also what many participants did not like, the bar chart did not excite them, and many admitted to entering the app and viewing their data less frequently while using the bar chart:

> "*It [the bar chart] is not as fun to look at. For one, I feel like I don't check it as often, because with the flower chart, you had it on the lock screen. And then you checked there and went into the app more often to look. Also, you don't think about going into the app as often when you're using the bar chart. And then I didn't think it was as interesting. Because it was a bit cooler when you could look at more specific flowers.*" ($P4_A$)

Not having the option to have the bar chart on their lock screen could be part of the

reason for this, but even participants who did not use the lock screen feature, also reported viewing their data less frequently while using the bar chart version. It could also be affected by which version the participants began with, as participants might become less interested in the app overall. There were more participants from group A, that began with the flower field, that thought the bar chart was less interesting than it was in group B, but more participants would be needed to see if this is a pattern or not.

7.2.3.2 Participants' thoughts on the flower field

In many ways, the appeal and critique of the flower field were the opposite to the bar chart. The flower field made many participants feel excited about looking at their data, and made them want to explore it. The visualization made the positive usage more apparent, and users explained that they could feel proud about using their phone in a good way, instead of being disappointed about how much they used their phone that day, because they could see what they used it for:

> "*Yeah, I mean, the flower field was new and exciting. It felt exciting and fun, and it had pretty colors, and it made you happy. I didn't think 'Oh, now I've spent so many hours on this.'*" ($P6_A$)

On the other hand, some also disliked seeing the wilted flowers, which then motivated them to use their phone better, and in some cases even to put down their phone and do something else. While many appreciated the aesthetics of the flower field, for some people it seemed that rather than inviting the user to explore more, it distracted them, making them focus more on how it looked rather than interacting with it. $P6_A$ describes:

> *"I think I got too distracted by how nice it looked, sort of. I don't think I'm reflecting differently, just that I got a bit more information.".*

Still, many participants also said that the flower field was informative, and gave them good insight into how they used their phone:

> "*The flower field is more informative. Even though it doesn't seem like it is as data driven, in theory, because like a bar chart is very, scientific and statistical, and it is supposed to be the simplest visualization for data, but the flower field, the different kinds of visualization, and the different kinds of flowers associated with your data, actually made it kind of easier to parse, and mentally make categories, that mapped my usage of my apps to, like the time spent on them.*" ($P7_B$)

Specifically, with the visualization focusing on the sessions rather than the total time, many users had realizations about how they used their phones for short moments, opening and closing apps frequently. Some also could reflect more on the correlation between category on session length, which was easier to see in the flower field rather than the bar chart.

There were however also participants who did not think that the flower field gave them a good overview of their data. This was partly because, as previously men-

tioned, the total time was not visible, but also that users got confused about flower sizes and categories. One user, $P2_A$ did not understand that the flower size correlated to the session's length (after it was explained in the interviews though $P2_A$ thought it made sense and was surprised they did not think of that), and three other users, $P4_A$, $P9_B$ and $P12_B$, did not at first understand that the categories were represented by different flowers. A consequence of having a flower for each session, even very short ones, was also that some users felt that their flower field looked full even though they hadn't used their phone for a large amount of time that day. This then led to a discrepancy, that they felt they had not used their phone very much, but the flower field did not reflect that. A few testers also complained about the contrast between some of the flowers and the background, especially in the brightest spot on the field.

7.2.4 Participants' thoughts on the lock screen wallpaper

Not all the participants had the flower field as their lock screen, $P1_A$, $P6_A$ and $P9_B$ did not, but for those who had, everybody had a positive experience with it. Many just appreciated it for the aesthetics, but for some, it also made them look at their data more, as it was there every time they picked up their phone. $P4_A$ also said that it became a conversation piece, and that those who used the app could compare their fields.

8

Discussion

This study set out to explore what design considerations are important when creating a digital tool that encourages reflection on smartphone use, with a focus on promoting meaningful rather than minimized usage. Two questions were in focus, what effects letting the user evaluate their phone use themselves had, as well as how visualizing this evaluation to the user would influence their reflection and behavior.

By letting the users themselves choose how they categorize their time, it resulted in the users feeling that their data was more representative of their actual behavior However, it does come at the cost of distracting and annoying the user with always having to make a choice. It was found that the categorization itself is important for reflection, with users thinking about the purpose of using an app in a specific moment. With many apps having multiple purposes, it allows users to distinguish between the usage that they enjoyed, and the usage they did not.

When comparing the experience of using the more conventional bar chart with the flower field visualization, eight participants preferred the flower field, with the main reason being that it felt more exciting and visually pleasing, while still presenting informative data. However, as three participants instead preferred the bar chart, it also became clear that keeping some familiarity in the visualization is important. Some participants misunderstood what different parts of the flower field meant, and appreciated the conventionality of the bar chart. In addition, some participants also preferred the bar chart because they felt it was more informative and easy to get an overview.

In the coming sections the effects the app had on reflection and behavioral change are discussed in more detail, as well as the trade-off between conventional and more novel visualizations. The discussion then shifts to reasons for different preferences of visualizations, problems with the flower field visualization and how the categorization increased reflection. Finally, ethical considerations, limitations, design implications and future work are presented.

8.1 Increase in reflection and behavior change

This study shows that users reflected more and became more aware of their smartphone use when presented with a more representative visualization of their self-perceived phone use. Increased reflection can lead to more awareness of when and

why they are using their phone, leading to a more mindful and conscious use of their phone [21]. This aligns with the METUX model [24] and the concepts of autonomy and competence, as the act of categorizing their smartphone use can support the need of autonomy as they can get an understanding of how well their smartphone use aligns with their values. Seeing their use visualized can also support the need of competence, as the user gets more insights into their behavior, and may get a better understanding of what can improve. A few participants even reported a change in behavior, which contradicts what some previous studies found [33], [36], [81], that simply visualizing the data does not lead to any change in behavior. However, as the scale of the study was quite small and had a rather short duration, further research would have to be done in order to draw any conclusions. Additionally, many participants did not report any change in behavior, which was somewhat to be expected as previous research indicates that interventions may be needed for behavior to change [33]. Furthermore, it can be difficult to distinguish whether it was the categorization itself, the visualization aspect, or both of them combined that lead to more reflections and potential changes in behavior. For example, one participant said that they felt sad when seeing their withered flowers, and therefore became motivated to use their phone better. Another participant said that they thought about whether they would categorize their current session as waste of time or procrastination even before the overlay was shown, and reflected on whether they should continue using the app. This suggests that both the categorizing part and the visualization part could be a factor in the potential behavior changing aspect.

8.2 Trade-off between conventional and new

Another factor to consider was how to approach the trade-off between the more conventional, but boring, bar chart and the more exciting, but sometimes confusing, flower field. With conventionality comes familiarity, and the user does not have to learn how to read and use the visualization, they already know from previous experiences, which can minimize confusion. However, it can also have drawbacks by not engaging the user and making them excited to explore their data. For areas where a task needs to be performed as efficiently as possible, familiar visualizations with good utility can be very important, but for a topic such as smartphone use, our research suggests that the more important task at hand could be to motivate the user to at all view or engage with their data, which can be achieved by focusing more on the attractiveness quality. However, finding the appropriate midpoint between conventionality and novelty has proved to be difficult and highly individual, as a new visualization can be confusing and some users can misunderstand how to interpret the visualization. Making the visualization exciting and fun can also go too far, with the user becoming distracted from the data at hand and making them forget that it is a visualization they can interact with. This is similar to the result from Wang, Tanahashi, Leaf, *et al.* [13], that abstract visualizations can help to intrigue users to explore their data, but that taking the abstractedness too far can become confusing.

However, this study shows the weight of exploring visualizations out of the ordinary,

and the impact it can have on the users inclination to view and explore data. While it is not in this study possible to draw any conclusions on how these aspects change over time, having more innovative visualizations is something that should be explored more when developing tools for smartphone use.

Another thing to note is that several of the participants said that they gained a more positive view of their smartphone use while using the app. Some became happy and proud, or were overall more satisfied with their use when they saw how much their smartphone use that was actually meaningful. As screen time often have been portrayed in media as harmful [2], many researchers have pointed out the importance of considering the complexity of screen use and the context in which it is used [7]. Simply showcasing the total number of hours a user has spent in a day can be problematic as it does not consider what the user has actually done, but rather makes the user feel guilty for spending much time on their phone. Lanette and Mazmanian [2] suggest that it is better to encourage curiousness and reflection about smartphone use which this study has tried to do. It has indicated that creating more nuanced and positive visualizations of screen time could be a part in helping users gain a better understanding and a more positive view of their screen time. This also aligns with views on digital well-being, especially with Vanden Abeele [17] dynamic model of digital well-being, with a focus on finding a balance between the benefits and drawbacks of technology. In this case, the Flower application can help users gain a better understanding of their smartphone use, allowing them to work on the use they classify as problematic, but also not feel bad about the necessary and enriching use.

8.3 Different insights from the visualizations

As the two visualizations were quite different, participants obtained different information out of them depending on which version they had. For example, in the flower field it was easy to see how frequently different apps were opened and closed each day. It was also easier to see a general overview over which categories were most present, and to see connections such as communication sessions tending to be smaller flowers (shorter duration) or procrastination sessions tending to be larger flowers.

The bar chart on the other hand, had more focus on the total amount of time each category was used, but it was harder to get an overview over the amount of sessions and their duration when they were not visible on one screen and had to be scrolled through.

It can be discussed whether both visualizations fulfilled all the Vitruvius triangle requirements, see figure 2.2. While no quantitative utility evaluation was performed, based on the interviews it can be concluded that both visualizations were useful and informative, but the bar chart was by many regarded as the more efficient visualization, quickly showing the total time for each category. However, the flower field visualization also allowed for easy overview, but from another perspective of the data. For the soundness criteria, the bar chart is a tried and true visualization

for countless different contexts and can be applied to most datasets. In comparison, the flower field visualization might be less applicable to other datasets, there is for example a limit to the amount of flowers that can be shown on the screen at once. However, the possibilities of showing complex data can be greater with the flower field visualization, as both type, size and placement can be coded to an attribute. Finally, since the attractiveness requirement were most in focus for the flower field visualization, it is stronger for this visualization. Many participants described the visualization as pretty and exciting, which showcases the beauty and novelty of the visualization. In contrast, the bar chart was instead often described as boring, partly because it is a very common and traditional visualization.

As the two visualizations differed significantly between how they visualized the data, many of the participants said they would like to see a combination of the two, or have a way to easily switch between them. This would be a good idea if the app were to be released on the market. We chose to have them be two separate views in order to be able to easily compare between them during the study. For example, it became clear that many participants looked at the data less frequently when they had the bar chart than when they had the flower field.

Since users also could have the flower field visualization on their lock screen, it served to remind the users to look at their data more. An interesting aspect of having the flower field on the lock screen is also that it becomes more public, as surrounding people might see it, and it can spark conversation. This element was something that only one person mentioned, but it would be interesting to explore further, and could be investigated if the duration of the test would be longer.

8.4 Preference of visualization

An interesting observation is how participants that used the flower field first preferred that visualization, or at least wanted a combination of both, but the participants from the group that started with the bar chart had more mixed opinions. The reason for this could be that people might have a bias towards the version that they get used to first. For example, if the user started with the flower field view, they might have explored it more at first as the app was new and exciting, and they could then have gotten used to that visualization. Then, when they changed version to the bar chart, it might have seemed more boring in contrast. However, if the user instead started with the bar chart, they would get more familiar with that visualization. Since the bar chart could be easier to understand, it might have been quite a shock when they changed to the flower field version that is more abstract and could then have been perceived as more confusing in comparison.

This is all speculation, however, and would need further study in order to be able to draw any conclusions. As the sort of visualization that people prefer can differ considerably depending on their personality and personal preferences [12], it might just be a coincidence that more of the participants that prefer a more conventional and data based visualization happened to be in the test group that started with the bar chart visualization.

Furthermore, in their study, Terzimehi, Haliburton, Greiner, *et al.* [21] also concluded that people's preferences are very personal and that an app that works well for one person might not work as well for another. This can be concluded from this project as well, as participants had different preferences of what visualization they liked. Additionally, they used the app very differently when it came to how often they opened the app or how many default categories they set.

8.5 Confusion around the flower field visualization

There were three participants, $P4_A$, $P9_B$ and $P12_B$, that did not understand that the flower types represented different categories in the beginning. They eventually understood, but it can be discussed whether the category to flower type mapping should be clearer. Perhaps the reason these participants did not understand at first was because they heavily used default categories and rarely got the overlay showcasing the flowers next to their respective category. It could also be an indication that they did not explore their data in the flower field as much. The confusion could also have risen from the fact that participants first got used to the bar chart view and the bar chart version of the app, and reflected less when later faced with the flower field view and the flower field version of the overlay.

In the introduction to the final tests we intentionally did not tell the participants too much about the flower visualization or that the different types of flower and flower sizes meant anything. The intention was that the participants themselves should explore the data and make their own findings and connections, which most of them did. However, this might also have led to some misunderstandings or misinterpretations of what the flowers represented. For example, one participant, $P2_A$ did not reflect at all that the flower sizes represented anything, they just thought it was random and therefore thought the flower field visualization of the data was not as representative as it might have been.

There were also a few participants that said they found it difficult to remember which flower represented which category. A solution to this problem and the problem that some participants did not understand that the flower types represented different categories, could be to also include an icon of the flowers in the drop down menu where the default categories are set. Perhaps the icon could also be included in the info box, or there could be a help page where this is explained.

It can also be discussed whether it would give better results if the participants were to be discouraged from using too many default categories during the introductory meeting. Perhaps it would have made the participants that heavily used them reflect more. The default categories could also be removed from the setup pages, to not encourage the user to set them immediately. However, as it turned out now, it was easier to see the difference between the participants that used default categories and those who did not, making it clearer what effect the overlay actually had on the users.

8.6 Categorization was important for reflection

Another key finding was that the process of categorizing sessions in the overlay was very important for the user to reflect and become more aware of what they used their phone for, which can be the first step to more mindful and meaningful phone use [19]. This study was a bit similar to MindPhone [21] that made the user reflect on what they were going to use the app for before they gained access to it. In their research they concluded that making the user more aware of what they were using their phone for, as well as their surrounding physical world, helped users to become more mindful of their phone use and to sometimes put it away in favor of doing something else. While our app did not ask the user anything before using an app, it showed somewhat similar results, that people reflected more during the actual use of the app, even when the overlay came after the app was closed.

While, similarly to this study, Finesse [35] and FinerMe [36] had users evaluate their phone use in an overlay after they used a social media app, they did not discuss this or research whether it got the users to reflect more. Instead, Finesse [35] researched more about user behaviour and what features of social media apps led to regretful behaviour, and FinerMe [36] focused more on the difference between feature-level and app-level interventions.

Something that these studies did not do either was to visualize this data to the user, how much and what usage was regretted in Finesse and FinerMe, and what was written in the prompts for MindPhone. Furthermore, while Finesse and FinerMe had the user evaluate their screen use, it was only done for a few specific social media apps via an experience sampling method, and it was only whether they regretted it or not. In this study, however, the choices were a bit more nuanced, having several different categories of what the app could be used for, which was very similar to what Lukoff, Yu, Kientz, *et al.* [18] did in their research [18]. Lukoff, Yu, Kientz, *et al.* [18], however, also used an experience sampling method and therefore did not acquire data for the users' smartphone use as a whole. They also did not visualize this back to the user [18]. In contrast, users in this study were able to evaluate and classify all the apps on their phone, gaining a more comprehensive overview of their entire phone use. Although, this came at the cost of either annoying the user with too many prompts or getting less accurate data with too many default categories.

8.7 Ethical considerations

The study did not involve vulnerable target groups, and we did not work with any sensitive data, such as health or similar. Still, smartphones are personal objects with personal information, which has been taken into account during the study.

Since the app developed tracked the amount of time participants spent on their phone, and which specific apps they used, restrictions were made on what type of data that was gathered. With the use of the accessibility service, a significant amount of information could be accessed, like text and content on the screen. However, we did only utilize a small portion of the possible events that could be tracked,

in order to track which app was currently in use and for how long. To be able to set default categories on apps, the Flower app also had access to all apps on the participant's device. However, the data was only saved locally on the participant's own devices, and we did not have any access to the data until the participants themselves sent the data during the last interview. The data that was sent to us was only a subset of the data the Flower app accessed, only the data needed for the app sessions, which was the app that was used, the duration and start time, and what category the user categorized it as, was sent. For example, this meant that we did not have access to all apps participants had on their device, but only the ones they had used during the testing period. Even though nobody withdrew from the test, it would have made it easier to do this, since they could choose to simply not send their data.

This approach to data collection is in line with other similar studies. Cho, Choi, Kim, *et al.* [35] did not collect data about the content on the screen, but only the information needed to determine which feature the user was using at each moment. Sathya and Nakagaki [52] also only tracked the data they need, in this case the titles of the Youtube videos the user watches, and deleted the data once it had been used. Boundaries like these are important so that participants feel comfortable using the tool and can have trust in that their data is handled with respect.

When gathering data through interviews, it is also important that it is stored securely and that anyone that has data gathered about them has given informed consent [61]. In many countries an informed consent means that the participants are informed about what the study is, what data will be collected, how it will be collected, and what it will be used for. It is also important that participants are informed that they can withdraw at any given moment and if they do so their previously gathered data will not be used [61]. During the introductory meeting, we explained to the participants the aim of the project, how the test would work and what was expected of them, as well as what data would be collected. They were also told that they could withdraw from the test at any point, and that their data would then not be used in the report. The interviews were recorded with the participants permission, and were transcribed either manually or with the AI tool whisper, which was run locally on our own computers. For thematic analysis, the coding tool Taguette [72] was used. This was self-hosted on our own server, avoiding the possibility of other parties having access to the data in any way. In this report, when using participants' smartphone data or excerpts from interviews, participants are anonymized, and instead referred to by a number.

8.8 Limitations

Even though the study garnered many interesting insights it did have some shortcomings. Unfortunately there were the bugs that made the logging of sessions not be entirely reliable. However, as it was a qualitative study and the insights gained from the user testing and interviews were in focus, this is more acceptable than if it would have been a quantitative study of the data. As these bugs were more device specific and did not appear on the phones we tested during development, it was

difficult to notice them before the actual user testing. In order to discover them we would have needed to test the app on a smaller group of users before the real user tests. However, as time was limited this was not really an option and it was hard to predict that it would have been needed.

Furthermore, the user test group was quite small, 12 participants in total and was a convenience sample (only university students). Due to technical limitations, only users with Android devices could participate in the study.

8.9 Design implications

Despite these limitations, this study brought about some important findings, that can be utilized when doing further research and design on this topic.

One implication is that, when designing with smartphone use as a design material, innovative and exciting visualizations are worth exploring, in order to make users more interested to explore their data. Although, caution has to be taken to balance the visualization between conventionality and abstractedness, in order to not confuse the user.

It also became clear that preferences of visualizations can be highly personal, and offering options or an ability to customize the visualizations seem to be a very important factor in order to accommodate different types of people.

Another consideration is that representative data is very important when visualizing smartphone use. Without this, users can feel a discrepancy if the data does not represent their actual behavior.

In addition, the act of making a choice on how an app session felt can be an important moment of reflection, although, there can be a fine balance between reflection and annoyance if the choice disrupts the user.

Finally, when designing around smartphone use, just as it is important to showcase the problematic use, it is also important to showcase the use that users felt good about.

8.10 Future work

While this study was able to investigate how an application like Flower can impact users everyday life, further research on long term effects of using similar applications is needed. The novelty of the app could be a strong contributing factor to increased reflections and changed behaviour. As many other screen time tools seem to be less effective over time, there are many aspects to consider that might not hold up over longer periods of use. For example, in the study, some participants indicated that the app could make them feel fatigued or annoyed of having to make decisions of how to categorize their use. This fatigue and annoyance may worsen over time, but it is also a possibility that users get used to it and that the categorization becomes an integrated part of app usage. We see a possibility of an intersection of

this research with AI, as an AI model might be able to learn how a user categorize apps depending on time, location and duration, which might decrease the amount of manual categorization that needs to be done. However, this would come with its own challenges and ethical considerations, and while it might decrease annoyance and fatigue, it would also take away possible moments of reflection.

Another aspect that was not explored in this study is one of personalization. One thing that became apparent in this study was how differently users used their phones, and how they thought about categories. Making it possible for users to choose which categories they want, and create their own, might allow users to make their data even more representative of how they actually use it. Additionally, adding customization on which flower or color corresponds to which category may also help users feel more connected to the visualization, and might make the visualizations less confusing, since the users are in charge of it.

This study showed that exploring alternative and more abstract visualizations is worthwhile. The flower field visualization created in this study is just an example of how smartphone use can be visualized, and there are countless more possibilities. Other alternate metaphors we chose between were for example; a city where each house is a different session with color depending on the category, or where sessions were represented as fruits in a fruit tree, or as fish in a fish tank. This also ties into the possibilities of personalization, where users could pick the visualization that they enjoy the most, maybe ranging from more conventional like the bar chart, to even more abstract versions than the one explored in this study. To support long term use, visualization that focus on larger spans than just a day could also be added. Showing a week, month, or year at once could allow for deeper reflection and comparison.

As many participants in this study enjoyed the aesthetic aspect of the flower field visualization, this is also a possible aspect to explore more. For this specific visualization, different lighting depending on time of day could be added, and animations to make the visualization feel more alive. Similar visualizations exist in other types of applications, for example in the weather app Yowindow [82].

A further area to research is also how an app like this would work if interventions were added as well, for example reminders of how much time has been spent in an app, or design friction interventions after an app has been used a set amount of time, like in FinerMe [36], see section 3.2.2. While the focus of this study lies on visualization and reflection, for those users who feel they use their phone in a problematic way, a combination of this more representative visualization in combination with interventions could be worth examining. As there has also been previous research on the impact of feature-level tracking [35], [36], it would be interesting to explore how smartphones use on a feature-level could be visualized.

As smartphone use is so nuanced and a single app can be used for multiple different purposes, there needs to be more options of visualizations that represents how users actually uses their phones. Further research needs to be done on how more creative and novel visualizations can be used to increase the users engagement and reflection. As there are many bad connotations associated with phone use, it could also be

important to reward the user when they actually use their phone in a meaningful way and not shaming them for using it normally.

9

Conclusion

This thesis set out to explore what should be considered when designing a tool for inspiring reflection about smartphone use, that puts an emphasis on meaningful smartphone use. This was done through creating an app, Flower, that allowed the user to themselves categorize how they used their phone after each app session and then visualized it back to the user. The app included two different visualizations in order to investigate how different visualizations might affect the result, one more conventional bar chart view and one more novel view of a flower field where each flower represented one app usage session. The app was then user tested with a within-subjects approach for eight days, where each participants tried the bar chart version for four days, and the flower field version for four days.

The study found that the overlay and categorization aspects of the app led to users reflecting more, and overall feel like the data that was visualized was more representative of their actual use. It was also found that most testers preferred the more novel flower field visualization, even though some found it confusing. Many appreciated the aesthetics of it, and it made them want to explore and look at the data more. However, there were also testers that appreciated the conventionality of the bar chart and liked that they could see the total amount of time spent on the different categories.

These results suggest that it is important to not view all smartphone use the same, but instead consider the different nuances of context and what different apps are actually used for. The study suggests that having the users categorize their screen-time themselves in the moment makes them reflect more on whether they enjoyed the time they spent or not. Furthermore, it revealed that having more novel and aesthetic visualization of data could increase the users interest in it and encourage them to explore it more. However it is important to consider the clarity of the data as a more abstract visualization can be confusing compared to more familiar and conventional visualizations.

The findings should be interpreted in light of the following limitations. There were some bugs that made the time logging sometime be inaccurate to varying degrees for different users. Furthermore, the sample size was not very large with 12 user testers in total, all of them university students.

Further research needs to be done on how more creative and novel visualizations can be used to increase the users engagement and reflection. It could also be further

researched how a study like this would hold up in the long run and with a larger quantity of testers. Additionally, it would be interesting to see if more personalized visualizations would increase users' self-reflection and understanding.

Overall, this thesis contributes to a deeper understanding of how more creative and novel visualizations help motivate users to interact with their data and how letting users categorize their use in the moment leads to more representative data and increased reflection.

Bibliography

[1] R. L. N. Lara N. Wolfers and N. Walter, "Too much screen time or too much guilt? how child screen time and parental screen guilt affect parental stress and relationship satisfaction," *Media Psychology*, vol. 28, no. 1, pp. 102–133, 2025. DOI: 10.1080/15213269.2024.2310839. eprint: https://doi.org/10.1080/15213269.2024.2310839. [Online]. Available: https://doi.org/10.1080/15213269.2024.2310839.

[2] S. Lanette and M. Mazmanian, "The smartphone "addiction" narrative is compelling, but largely unfounded," in *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI EA '18, Montreal QC, Canada: Association for Computing Machinery, 2018, pp. 1–6, ISBN: 9781450356213. DOI: 10.1145/3170427.3188584. [Online]. Available: https://doi.org/10.1145/3170427.3188584.

[3] A. Orben, "Digital diet: A 21st century approach to understanding digital technologies and development," *Infant and Child Development*, vol. 31, no. 1, e2228, 2021. DOI: https://doi.org/10.1002/icd.2228. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/icd.2228. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/icd.2228.

[4] Folkhälsomyndigheten, "Digitala medier och barns och ungas hälsa - en kunskapssammanställning," Folkhälsmyndigheten and Mediemyndigheten, Tech. Rep., 2024. [Online]. Available: https://www.folkhalsomyndigheten.se/contentassets/20a0ad3202d54bc9be156ff3e407b55c/digitala-medier-barns-ungas-halsa-kunskapssammanstallning.pdf.

[5] N. Stiglic and R. Viner, "Effects of screentime on the health and well-being of children and adolescents: A systematic review of reviews," *BMJ Open*, 2019. DOI: 10.1136/bmjopen-2018-023191.

[6] L. Kaye, A. Orben, D. Ellis, S. Hunter, and S. Houghton, "The conceptual and methodological mayhem of screen time," *International Journal of Environmental Research and Public Health*, vol. 17, p. 3661, May 2020. DOI: 10.3390/ijerph17103661.

[7] J. Zosh, M. Jing, J. Shawcroft, and A. Starks, "Moving beyond quality and quantity: Approaching childrens digital media use with more nuance," *Translational Issues in Psychological Science*, vol. 9, pp. 181–185, Sep. 2023. DOI: 10.1037/tps0000378.

[8] S. Livingstone, "The rise and fall of screen time," en, in *Masters of Media: Controversies and Solutions: Controversies and Solutions, Volume 1*, V. C. Strasburger, Ed., vol. 1, London, UK: Rowman and Littlefield, Jul. 2021, p. 16.

[9] R. S. Al-Mansoori, D. Al-Thani, and R. Ali, "Designing for digital wellbeing: From theory to practice a scoping review," *Human Behavior and Emerging Technologies*, vol. 2023, no. 1, p. 9 924 029, 2023. DOI: https://doi.org/10.1155/2023/9924029. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1155/2023/9924029. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1155/2023/9924029.

[10] O. Sosa-Tzec, "Delightful screen time visualizations: A sketching-based exploration," *Nordes 2023*, vol. 1, Jun. 2023, ISSN: 1604-9705. DOI: 10.21606/nordes.2023.133. [Online]. Available: http://dx.doi.org/10.21606/nordes.2023.133.

[11] D. Huang, M. Tory, B. A. Aseniero, *et al.*, "Personal visualization and personal visual analytics," en, *IEEE Trans. Vis. Comput. Graph.*, vol. 21, no. 3, pp. 420–433, Mar. 2015.

[12] H. Schneider, K. Schauer, C. Stachl, and A. Butz, "Your data, your vis: Personalizing personal data visualizations," in *Human-Computer Interaction – INTERACT 2017*, R. Bernhaupt, G. Dalvi, A. Joshi, D. K. Balkrishan, J. O'Neill, and M. Winckler, Eds., Cham: Springer International Publishing, 2017, pp. 374–392, ISBN: 978-3-319-67687-6.

[13] S. Wang, Y. Tanahashi, N. Leaf, and K.-L. Ma, "Design and effects of personal visualizations," *IEEE Computer Graphics and Applications*, vol. 35, no. 4, pp. 82–93, 2015. DOI: 10.1109/MCG.2015.74.

[14] S. Csibi, M. D. Griffiths, Z. Demetrovics, and A. Szabo, "Analysis of problematic smartphone use across different age groups within the components model of addiction," *International Journal of Mental Health and Addiction*, vol. 19, no. 3, pp. 616–631, 2021.

[15] R. S. Al-Mansoori, D. Al-Thani, and R. Ali, "Designing for digital wellbeing: From theory to practice a scoping review," *Human Behavior and Emerging Technologies*, vol. 2023, no. 1, p. 9 924 029, 2023. DOI: https://doi.org/10.1155/2023/9924029. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1155/2023/9924029. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1155/2023/9924029.

[16] R. A. Calvo and D. Peters, *Positive computing: technology for wellbeing and human potential*. MIT press, 2014.

[17] M. M. Vanden Abeele, "Digital wellbeing as a dynamic construct," *Communication Theory*, vol. 31, no. 4, pp. 932–955, 2021.

[18] K. Lukoff, C. Yu, J. Kientz, and A. Hiniker, "What makes smartphone use meaningful or meaningless?" *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, no. 1, Mar. 2018. DOI: 10.1145/3191754. [Online]. Available: https://doi.org/10.1145/3191754.

[19] A. Hiniker, S.-S. Hong, T. Kohno, and J. A. Kientz, "Mytime: Designing and evaluating an intervention for smartphone non-use," *International Conference on Human Factors in Computing Systems*, 2016. DOI: 10.1145/2858036.2858403.

[20] N. Terzimehi, R. Häuslschmid, H. Hussmann, and m. schraefel m.c., "A Review & Analysis of Mindfulness Research in HCI: Framing Current Lines of Research and Future Opportunities," in *Proceedings of the 2019 CHI Con-*

ference on Human Factors in Computing Systems, ser. CHI '19, Glasgow, Scotland Uk: Association for Computing Machinery, 2019, pp. 1–13, ISBN: 9781450359702. DOI: 10.1145/3290605.3300687. [Online]. Available: https://doi.org/10.1145/3290605.3300687.

[21] N. Terzimehi, L. Haliburton, P. Greiner, A. Schmidt, H. Hussmann, and V. Mäkelä, "Mindphone: Mindful reflection at unlock can reduce absentminded smartphone use," in *Proceedings of the 2022 ACM Designing Interactive Systems Conference*, ser. DIS '22, Virtual Event, Australia: Association for Computing Machinery, 2022, pp. 1818–1830, ISBN: 9781450393584. DOI: 10.1145/3532106.3533575. [Online]. Available: https://doi.org/10.1145/3532106.3533575.

[22] J. A. Tran, K. S. Yang, K. Davis, and A. Hiniker, "Modeling the engagement-disengagement cycle of compulsive phone use," *International Conference on Human Factors in Computing Systems*, 2019. DOI: 10.1145/3290605.3300542.

[23] R. Ryan, "Self determination theory and well being," *Social Psychology*, vol. 84, no. 822, p. 848, 2009.

[24] D. Peters, R. A. Calvo, and R. M. Ryan, "Designing for motivation, engagement and wellbeing in digital experience," *Frontiers in Psychology*, vol. Volume 9 - 2018, 2018, ISSN: 1664-1078. DOI: 10.3389/fpsyg.2018.00797. [Online]. Available: https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2018.00797.

[25] C. Chen, "Information visualization," *WIREs Computational Statistics*, vol. 2, no. 4, pp. 387–403, 2010. DOI: https://doi.org/10.1002/wics.89. eprint: https://wires.onlinelibrary.wiley.com/doi/pdf/10.1002/wics.89. [Online]. Available: https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wics.89.

[26] A. V. Moere and H. Purchase, "On the role of design in information visualization," *Information Visualization*, vol. 10, no. 4, pp. 356–371, 2011. DOI: 10.1177/1473871611415996. eprint: https://doi.org/10.1177/1473871611415996. [Online]. Available: https://doi.org/10.1177/1473871611415996.

[27] D. Cox, "Metaphoric mappings: The art of visualization," in *Aesthetic Computing*, P. A. Fishwick, Ed., MIT Press, 2006, pp. 89–114.

[28] E. Hoffman, *Time*, English. London: Profile books LTD, 2011, ISBN: 9781846680434; 9781847652591.

[29] O. Turel and D. R. Cavagnaro, "Effect of abstinence from social media on time perception: Differences between low- and at-risk for social media "addiction" groups," *Psychiatric Quarterly*, vol. 90, no. 1, pp. 217–227, Mar. 2019, ISSN: 1573-6709. DOI: 10.1007/s11126-018-9614-3. [Online]. Available: https://doi.org/10.1007/s11126-018-9614-3.

[30] P. Verduyn, N. Gugushvili, and E. Kross, "Do social networking sites influence well-being? the extended active-passive model," *Current Directions in Psychological Science*, vol. 31, no. 1, pp. 62–68, 2022. DOI: 10.1177/09637214211053637. eprint: https://doi.org/10.1177/09637214211053637. [Online]. Available: https://doi.org/10.1177/09637214211053637.

[31] P. Verduyn, M. R. Oscar Ybarra, J. Jonides, and E. Kross, "Do social network sites enhance or undermine subjective well-being? a critical review.," *Social*

Issues and Policy Review, vol. 11, pp. 274–302, 2017. DOI: https://doi.org/10.1111/sipr.12033.

[32] J. A. Tran, K. S. Yang, K. Davis, and A. Hiniker, "Modeling the engagement-disengagement cycle of compulsive phone use," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI '19, Glasgow, Scotland Uk: Association for Computing Machinery, 2019, pp. 1–14, ISBN: 9781450359702. DOI: 10.1145/3290605.3300542. [Online]. Available: https://doi.org/10.1145/3290605.3300542.

[33] F. Rahmillah, A. Tariq, M. King, and O. Oviedo-Trespalacios, "Evaluating the effectiveness of apps designed to reduce mobile phone use and prevent maladaptive mobile phone use: Multimethod study," *Journal of medical Internet research*, vol. 25, Aug. 2023. DOI: 10.2196/42541.

[34] F. I. Rahmillah, A. Tariq, M. King, and O. Oviedo-Trespalacios, "Is distraction on the road associated with maladaptive mobile phone use? a systematic review," *Accident Analysis & Prevention*, vol. 181, p. 106 900, 2023, ISSN: 0001-4575. DOI: https://doi.org/10.1016/j.aap.2022.106900. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0001457522003359.

[35] H. Cho, D. Choi, D. Kim, W. J. Kang, E. K. Choe, and S.-J. Lee, "Reflect, not regret: Understanding regretful smartphone use with app feature-level analysis," *Proc. ACM Hum.-Comput. Interact.*, vol. 5, no. CSCW2, Oct. 2021. DOI: 10.1145/3479600. [Online]. Available: https://doi.org/10.1145/3479600.

[36] A. Orzikulova, H. Cho, H.-Y. Chung, H. Hong, U. Lee, and S.-j. Lee, "Finerme: Examining app-level and feature-level interventions to regulate mobile social media use," *null*, 2023. DOI: 10.1145/3610065.

[37] Apple. "Use screen time on your iphone or ipad." (2024), [Online]. Available: https://support.apple.com/en-us/108806 (visited on 01/30/2025).

[38] Forest. "Forest." (), [Online]. Available: https://www.forestapp.cc/ (visited on 01/30/2025).

[39] Android. "Nya sätt att hitta en balans för dig och din familj." (), [Online]. Available: https://www.android.com/digital-wellbeing/ (visited on 01/30/2025).

[40] Antisocial. "Antisocial." (), [Online]. Available: http://antisocial.io/ (visited on 01/30/2025).

[41] ScreenZen. "Screenzen." (), [Online]. Available: https://www.screenzen.co/ (visited on 01/30/2025).

[42] one sec. "One sec." (), [Online]. Available: https://one-sec.app/ (visited on 01/30/2025).

[43] clearspace. "Clearspace." (), [Online]. Available: https://www.getclearspace.com/ (visited on 01/30/2025).

[44] D. J. Grüning, F. Riedel, and P. Lorenz-Spreen, "Directing smartphone use through the self-nudge app one sec," *Proceedings of the National Academy of Sciences*, vol. 120, no. 8, e2213114120, 2023. DOI: 10.1073/pnas.2213114120. eprint: https://www.pnas.org/doi/pdf/10.1073/pnas.2213114120.

[Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.2213114120.

[45] StayFree. "Stayfree." (), [Online]. Available: https://stayfreeapps.com/ (visited on 01/30/2025).

[46] Flora. "Flora." (), [Online]. Available: https://flora.appfinca.com/ (visited on 01/30/2025).

[47] Focus Plant. "Focus plant." (), [Online]. Available: https://play.google.com/store/apps/details?id=com.shikudo.focus.google&hl=sv (visited on 01/30/2025).

[48] androidcentral. "Google's digital wellbeing initiative: Everything you need to know." (2018), [Online]. Available: www.androidcentral.com/googles-digital-wellbeing (visited on 01/30/2025).

[49] The Verge. "How to use apples new screen time and app limits features in ios 12, apple is making it easier than ever to cut back on app overload." (2018), [Online]. Available: www.theverge.com/2018/9/17/17870126/ios-12-screen-time-app-limits-downtime-features-how-to-use (visited on 01/30/2025).

[50] Samsung. "What is digital wellbeing feature on samsung smartphone?" (2022), [Online]. Available: https://www.samsung.com/sg/support/mobile-devices/what-is-digital-wellbeing-feature-on-samsung-smartphones/ (visited on 01/30/2025).

[51] Huawei. "[digital balance] bring more balance into your life." (2021), [Online]. Available: https://consumer.huawei.com/en/community/details/topicId-158991/ (visited on 01/30/2025).

[52] A. Sathya and K. Nakagaki, "Attention receipts: Utilizing the materiality of receipts to improve screen-time reflection on youtube," in *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, ser. CHI '24, Honolulu, HI, USA: Association for Computing Machinery, 2024, ISBN: 9798400703300. DOI: 10.1145/3613904.3642505. [Online]. Available: https://doi.org/10.1145/3613904.3642505.

[53] H. W. J. Rittel and M. M. Webber, "Dilemmas in a general theory of planning," *Policy Sciences*, vol. 4, no. 2, pp. 155–169, Jun. 1973, ISSN: 1573-0891. DOI: 10.1007/BF01405730. [Online]. Available: https://doi.org/10.1007/BF01405730.

[54] R. Buchanan, "Wicked problems in design thinking," *Design Issues*, vol. 8, no. 2, pp. 5–21, 1992, ISSN: 07479360, 15314790. [Online]. Available: http://www.jstor.org/stable/1511637 (visited on 02/06/2025).

[55] J. Zimmerman, E. Stolterman, and J. Forlizzi, "An analysis and critique of research through design: Towards a formalization of a research approach," in *Proceedings of the 8th ACM Conference on Designing Interactive Systems*, ser. DIS '10, Aarhus, Denmark: Association for Computing Machinery, 2010, pp. 310–319, ISBN: 9781450301039. DOI: 10.1145/1858171.1858228. [Online]. Available: https://doi.org/10.1145/1858171.1858228.

[56] C. Frayling, "Research in art and design. " royal college of art," *Royal College of Art*, vol. 1, no. 1, 1993.

[57] W. Gaver, "What should we expect from research through design?" In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '12, Austin, Texas, USA: Association for Computing Machinery, 2012, pp. 937–946, ISBN: 9781450310154. DOI: 10.1145/2207676.2208538. [Online]. Available: https://doi.org/10.1145/2207676.2208538.

[58] M. Kochanowska, W. R. Gagliardi, and with reference to Jonathan Ball, "The double diamond model: In pursuit of simplicity and flexibility," in *Perspectives on Design II: Research, Education and Practice*, D. Raposo, J. Neves, and J. Silva, Eds. Cham: Springer International Publishing, 2022, pp. 19–32, ISBN: 978-3-030-79879-6. DOI: 10.1007/978-3-030-79879-6_2. [Online]. Available: https://doi.org/10.1007/978-3-030-79879-6_2.

[59] Design Council. "Framework for inovation." (), [Online]. Available: https://www.designcouncil.org.uk/our-resources/framework-for-innovation/ (visited on 02/07/2025).

[60] B. Hanington and B. Martin, *Universal methods of design : 125 ways to research complex problems, develop innovative ideas, and design effective solutions*, English, Expanded and revised edition. Beverly, MA: Rockport Publishers, 2019, ISBN: 9781631597480; 1631597485.

[61] Y. Rogers, H. Sharp, and J. Preece, *Interaction Design: Beyond Human-Computer Interaction*. Wiley, 2023, ISBN: 9781119901112. [Online]. Available: https://books.google.se/books?id=E26yEAAAQBAJ.

[62] T. Kelley and J. Littman, *The Art of Innovation*. Doubleday, 2000.

[63] Google. "Crazy 8's." (), [Online]. Available: https://designsprintkit.withgoogle.com/methodology/phase3-sketch/crazy-8s (visited on 02/07/2025).

[64] K. N. Truong, G. R. Hayes, and G. D. Abowd, "Storyboarding: An empirical determination of best practices and effective guidelines," in *Proceedings of the 6th Conference on Designing Interactive Systems*, ser. DIS '06, University Park, PA, USA: Association for Computing Machinery, 2006, pp. 12–21, ISBN: 1595933670. DOI: 10.1145/1142405.1142410. [Online]. Available: https://doi.org/10.1145/1142405.1142410.

[65] Agile Business Consortium, "Dsdm agile project framework handbook," in Agile Business Consortium, 2014, ch. MoSCoW Prioritisation, Accessed May 12, 2025. [Online]. Available: https://www.agilebusiness.org/dsdm-project-framework.html.

[66] D. Benyon, P. Turner, and S. Turner, *Designing interactive systems: People, activities, contexts, technologies*. Pearson Education Limited, 2005.

[67] Y.-K. Lim, E. Stolterman, and J. Tenenberg, "The anatomy of prototypes: Prototypes as filters, prototypes as manifestations of design ideas," *ACM Trans. Comput.-Hum. Interact.*, vol. 15, no. 2, Jul. 2008, ISSN: 1073-0516. DOI: 10.1145/1375761.1375762. [Online]. Available: https://doi.org/10.1145/1375761.1375762.

[68] N. van Berkel, D. Ferreira, and V. Kostakos, "The experience sampling method on mobile devices," *ACM Comput. Surv.*, vol. 50, no. 6, Dec. 2017, ISSN: 0360-0300. DOI: 10.1145/3123988. [Online]. Available: https://doi.org/10.1145/3123988.

[69] N. N. Group. "Between-subjects vs. within-subjects study design." (2023), [Online]. Available: https://www.nngroup.com/articles/between-within-subjects/ (visited on 05/05/2025).
[70] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, "Robust speech recognition via large-scale weak supervision," in *Proceedings of the 40th International Conference on Machine Learning*, ser. ICML'23, Honolulu, Hawaii, USA: JMLR.org, 2023.
[71] Purfview, *Whisper-standalone-win*. [Online]. Available: https://github.com/Purfview/whisper-standalone-win (visited on 02/19/2025).
[72] R. Rampin and V. Rampin, *Taguette*. [Online]. Available: https://www.taguette.org/.
[73] Kotlin. "Kotlin for android." (2024), [Online]. Available: https://kotlinlang.org/docs/android-overview.html (visited on 04/22/2025).
[74] Developers. "Build better apps faster with jetpack compose." (), [Online]. Available: https://developer.android.com/compose (visited on 04/22/2025).
[75] *Github*. [Online]. Available: https://github.com/ (visited on 04/28/2025).
[76] A. Developers. "Save data in a local database using room." (), [Online]. Available: https://developer.android.com/training/data-storage/room (visited on 04/28/2025).
[77] *Accessibility service api*. [Online]. Available: https://developer.android.com/reference/android/accessibilityservice/AccessibilityService (visited on 05/13/2025).
[78] *Procreate*. [Online]. Available: https://procreate.com/ (visited on 04/28/2025).
[79] V. Analytics. "Most popular mobile social networking apps in the united states as of september 2019, by average session duration." (), [Online]. Available: https://www.statista.com/statistics/579411/top-us-social-networking-apps-ranked-by-session-length/ (visited on 04/28/2025).
[80] *Figjam*. [Online]. Available: https://www.figma.com/figjam/ (visited on 05/14/2025).
[81] L. Zimmermann, "your screen-time app is keeping track: Consumers are happy to monitor but unlikely to reduce smartphone usage," *Journal of the Association for Consumer Research*, vol. 6, no. 3, pp. 377–382, 2021. DOI: 10.1086/714365. eprint: https://doi.org/10.1086/714365. [Online]. Available: https://doi.org/10.1086/714365.
[82] *Yowindow*. [Online]. Available: https://yowindow.com/ (visited on 05/09/2025).

A

Appendix 1

A.1 Prestudy interview questions

Introductory Questions

1. Is it okay if we record the interview and use parts of your answers in our report? Only the two of us will listen to the recording, and it will be deleted as soon as the project is finished. Ill be asking the questions, and (Stina/Hanna) will take some notes.
2. How old are you?
3. What is your main occupation?

Screen Time Questions

1. Would you like to show your screen time data from the past week? We wont show it in our reportits just to help reflect on it. Is it okay if we take a photo when you show it?
2. How do you feel about your screen time?
3. What do you spend the most time on?
4. Do you actively try to reduce your screen time?
5. Do you often check how much time youve spent?
6. Do you use any tools to limit your screen time?
7. Which ones? Do you feel they help? (ask them to show)
8. What do you do on your phone that you feel benefits you?
9. What do you do on your phone that you feel doesnt benefit you?
10. In what situations do you use your phone the most?
11. Under what circumstances do you tend to regret using your phone?

App Use Questions

1. Would you like to show us an app you use a lot? Please show us how and when you usually use it, what features you go through, and in what order.
2. What do you like about that app?
3. What do you dislike about that app?

A.2 Persona scenarios

A.2.1 Linda

Positive experience Linda is sitting at school in the lunch room and scrolls Instagram while she waits for her next lecture to start. She watches a reel of a girls outfit of the day, and she likes how it looks. She thinks about her own clothes that she has, and realizes she has almost the same pieces, but hadnt really thought about putting them together before. She saves the reel in the app, and then continues scrolling until her break has ended. Later, the next morning, she goes into her saved posts, and watches the reel again. She picks out the pieces from her wardrobe and puts them on, and feels happy about the new outfit she got inspired to put on.

Negative experience Linda comes home after a long day at school, and she feels tired and does not feel like doing anything yet. She sits down on the sofa and starts scrolling on TikTok. After fifteen minutes of scrolling she gets a reminder from her app timer saying she has five minutes left on her timer. She dismisses it and continues scrolling. Once five minutes has passed, she gets locked out of the app, but since she was in the middle of watching a video she clicks on the button that extends her time a few minutes. After a while, her partner comes home, and she feels guilty about how much time she has spent on her phone that afternoon.

A.2.2 Thomas

Positive experience Thomas decides to scroll on Pinterest to unwind after he gets home from work. There he finds a few interesting recipes he would like to try out. He feels inspired, pins them and decides to try out one of them the next day. He writes up the ingredients hes missing on his shopping list and tells his roommate about it and ask if he wants to try some too.

Negative experience His friend sends a funny TikTok to him on messenger, but since he doesnt have the app he cant watch it. He often feels a bit out of the loop, since he barely uses social media, therefore decides to install TikTok again to watch what his friend sent him. He watches it and then scrolls some more. He doesnt uninstall the app right away, instead he finds himself using it more and scrolls on it several times a day. He does this until he feels frustrated at himself for scrolling too much and uninstalls the app again.

A.2.3 Rose

Positive experience When she goes out on a walk she decides to put away her phone in her pocket on no disturb. She feels like she is more in the moment and is able to more fully immerse herself in her hike when she knows she wont get any notifications.

Negative experience Sometimes during lectures she feels herself wanting to pick up her phone to check if she got any new messages or check her email. She feels like she becomes distracted from her lectures or school work just thinking about it.

A.3 Global themes from pre-study with corresponding organizing themes

A.3.1 Designed for engagement

Stuck scrolling

When users feels stuck scrolling on endless feeds on social media or opening one thing that leads to another and then another. Often leads to a feeling of overuse and a negative viewpoint on screen time. Often happens after school/work or before bed.

Multiple purposes

When app designers include many different features in an app it makes the user unable to remove the app as they want some of the features, but might overuse some of the other, wishing they could just remove them.

App design

Malicious app design intended to make the user use the app as much as possible. For example, mixing the features/content the user actually wants with the content they usually get stuck scrolling on or making sure the app opens on the content that users usually get stuck on.

Feeling finished

Many participants explained how they usually check all new things first, notifications, new messages and updates. Additionally, they praised apps that naturally had an end, when there were no new things to watch, they exited the app. On the contrary, for social media apps where the content never ends, some participants expressed how they wished it did.

A.3.2 Compare to others

Sub-theme same as global theme, see 6.1.3.

A.3.3 Hard to resist

Compulsive use

When users check their phone for new updates excessively, or picks up their phone automatically, without thinking about why.

Always there

The smartphone is always a possibility, every time users are bored, awkward or dont have anything else to do. This is not always a bad thing, it can be convenient to be

able to read books on the phone, and it can be a good break during work to scroll for a little while, as it is always close by.

Distracting from the moment

Since it is always there, it can be difficult to resist to use at some times where it would probably be better not to. During social settings, or when there are other things to do, the smartphone can be distracting. Some participants mentioned how they actively put away their phones, removed notifications, or put it in airplane mode, to make it less distracting.

Sleep

Many participants used their phone right before bed, and many said that they knew this was bad for sleep, and could affects their sleep, but they still did it. Some did also say that using their phone made it easier for them to sleep.

A.3.4 Conscious of screen time

Reduction

This theme encapsulates how many participants were aware of their smartphone use, and the methods they used for reducing it. Getting notifications, using app timers, setting goals, increasing the friction to access the app, or just simply practicing self control when they felt they used their phone too much.

Bad visualization

Sometimes participants expressed how the time visualization was either a bit confusing to use and find, or that it did not match their perception. Since having a lot of screen time can feel bad for many, when they felt that what they did on their phone was useful or fun, but lead to a lot of screen time, it could still make them feel bad about their time.

A.3.5 Enriches life

Connection

Smartphones enables people to stay connected with each other, they can use communication apps to message or talk to each other, social media to share their lives and keep up to date with friends lives or share funny videos with each other.

Inspiring

A lot of participants use their phone to gather information and learn new things or to gain inspiration for and learn more about their hobbies and interests.

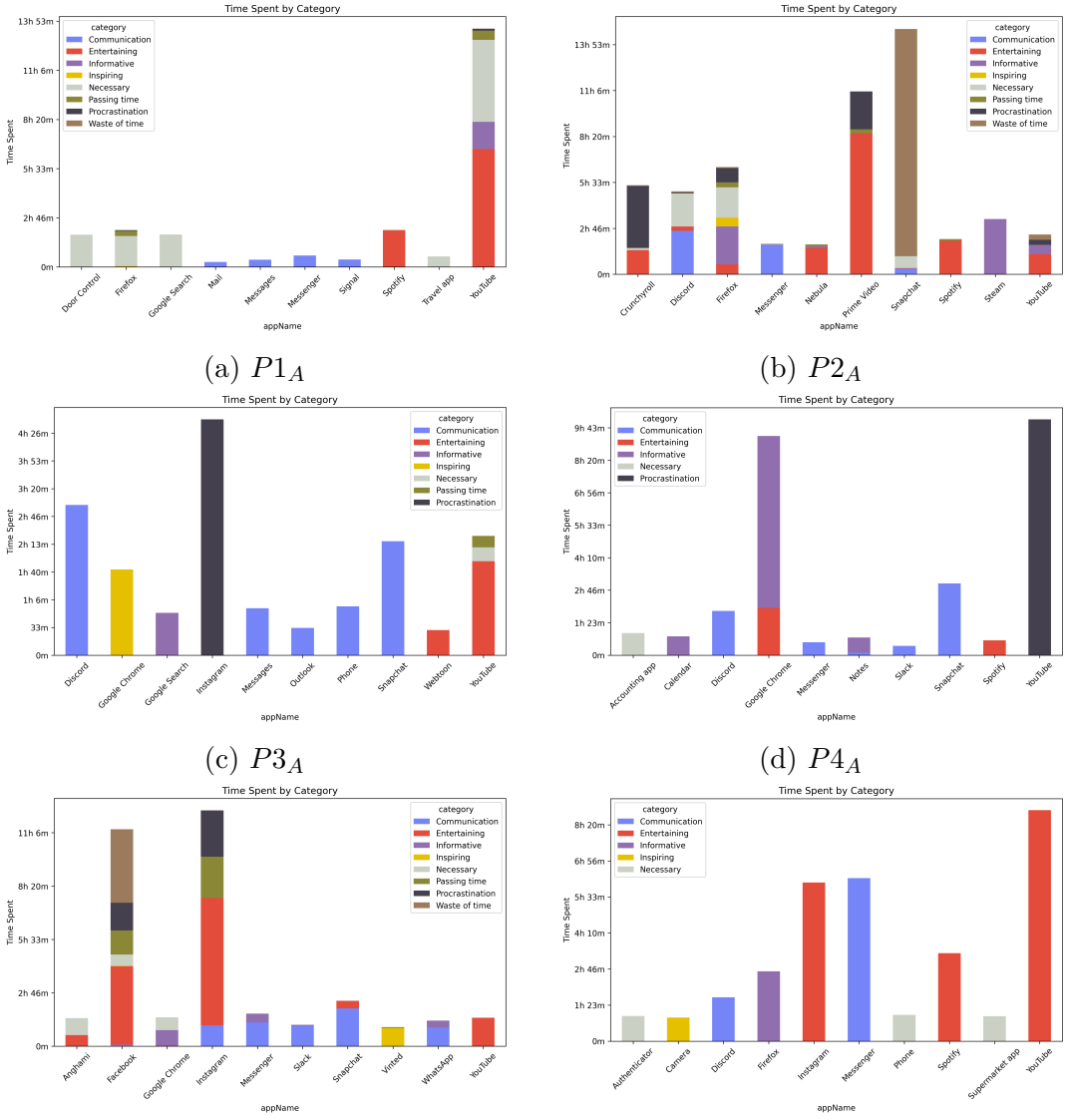
Necessary

Participants expressed how they used their phone for more necessary things like communication, seeking information, managing their economy, traveling or keeping track of news.

A.4 Final interview questions

1. Is it okay if we record the interview and use parts of your answers in our report? Only the two of us will listen to the recording afterwards, and it will be deleted as soon as were done with the project. I will be asking the questions, and (Stina/Hanna) will take some notes.
2. How old are you?
3. What has your experience of using the app been like? Feel free to compare the two versions of the app.
4. How do you feel when looking at your data now with the flower field compared to when you had the bar charts?
5. Which version of the app do you prefer? Why?
6. Have your mobile habits changed in any way? Did the data visualization affect this in any way?
7. Do you reflect differently on your mobile use compared to before you had the app? Did you reflect differently depending on which visualization you had?
8. Is it possible to get an overview of your mobile use? Why or why not? Does it differ based on the visualization?
9. What did you think about the categories? Was any category missing, or was any unnecessary?
10. How much did you use the preset categories? What was your thinking process when assigning them?
11. Which types of apps were always in the same category, and which ones varied?
12. Did you find it difficult to decide which category to assign to a session? Please give an example of when it was easy and one when it was difficult.
13. Did you gain anything from having the flower field on your lock screen (if you used it)? If so, what?
14. Did you feel the flowers reflected the categories well?
15. Is there anything else that could be improved now that youve tried the second version of the app? Is there anything youre missing?
16. How often did bugs occur? Did you delete those sessions, or are they still there?

A.5 Categorization of the testers top ten most used apps

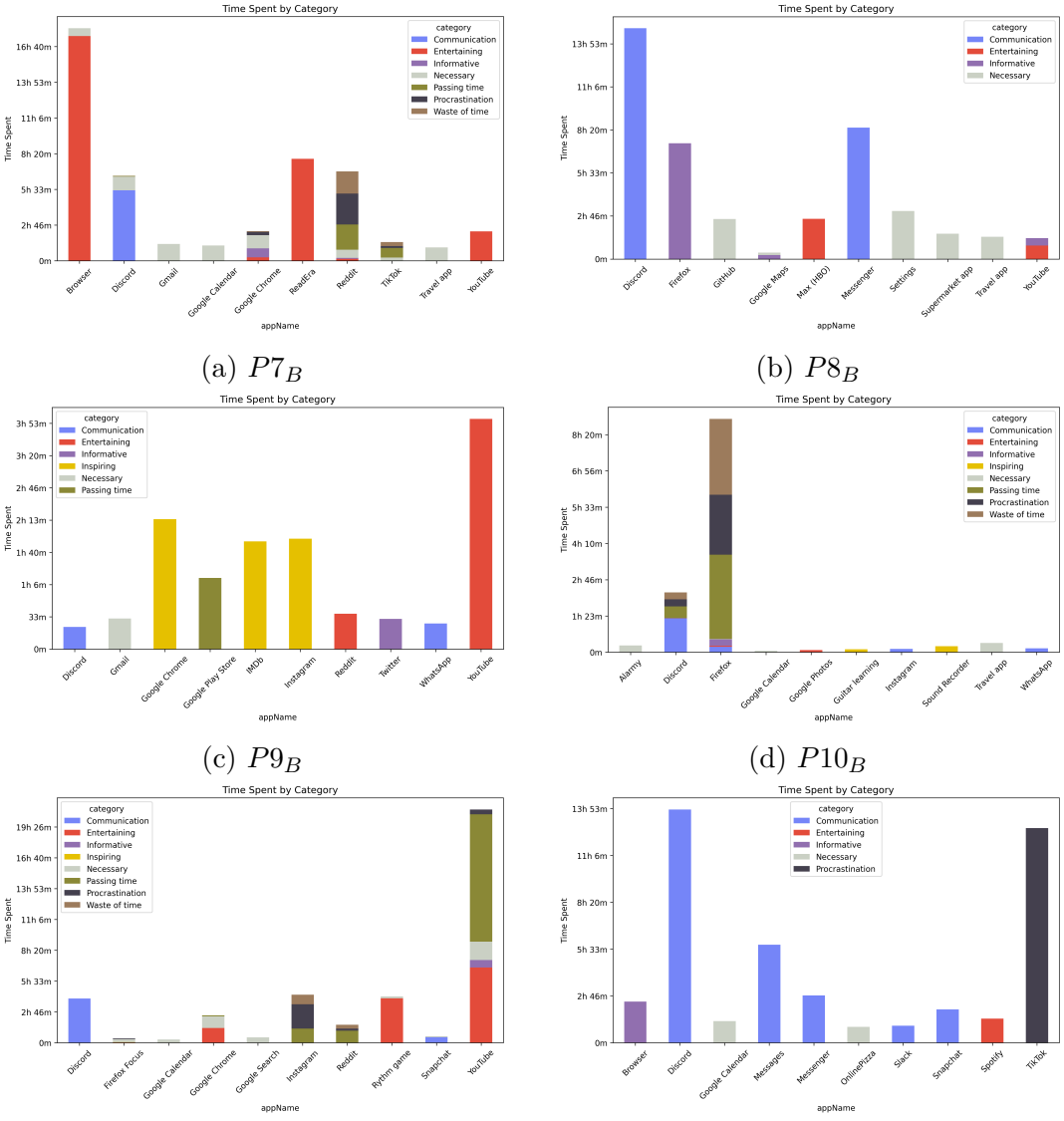


(a) $P1_A$ (b) $P2_A$

(c) $P3_A$ (d) $P4_A$

(e) $P5_A$ (f) $P6_A$

Figure A.1: Categorization of the testers top 10 most used apps, test group A



(a) $P7_B$ (b) $P8_B$

(c) $P9_B$ (d) $P10_B$

(e) $P11_B$ (f) $P12_B$

Figure A.2: Categorization of the testers top 10 most used apps, test group B