

OUTSMART YOUR PHONE

HOW COMMITMENT CONTRACTS CAN HELP US SPEND LESS TIME ON OUR
PHONES AND WHY THEY WORK BETTER FOR YOU THAN FOR ME

Max Valentin Bartels, 41193

Word count: 19572

Presentation: May 2019

Supervisor: Assistant Professor *Wiley Wakeman* - Stockholm School of Economics

Thesis coordinators: Associate professor *Susanne Sweet* and associate professor *Markus Kallifatides* – Stockholm School of Economics

Abstract

This master thesis sought out to address arising problems from increases in smartphone usage. Many of us spend multiple hours a day on our phones. While the benefits mobile devices can bring are obvious for most, our increasing usage of and psychological dependency on such devices come with serious side-effects. These negative impacts have attracted rising attention from media, practitioners, researchers, and individuals alike. Many people state they spend too much time on their phones but don't know how to achieve their goals of decreasing smartphone usage. To address this modern-day issue of self-control, this study draws on the established concept of commitment contracts. An experiment was conducted in which participants set out to decrease their smartphone usage by 30% for one week. Commitment contracts, designed so that participants had to perform a good deed for every day they failed to stay below a screen time limit, served as treatment for half of the participants. Compared to participants who only set themselves the goal to decrease smartphone usage, participants exposed to commitment contracts showed to be significantly better at achieving their new daily screen time goals. While results further suggest that relative smartphone usage reduction in percent could not be significantly improved through commitment contracts, the combination of these two outcome measures provides partial support for the effectiveness of commitment contracts in the context of smartphone usage. To investigate whether commitment contracts work better for some individuals than for others, this study also analyzed the relationship between commitment contract effectiveness and individuals' future-self-continuity – a measure of how similar and connected we feel to our future-selves. Results could not support the hypothesis that commitment contracts work better for individuals with low future-self-continuity than for individuals with high future-self-continuity. This study adds to the existing body of research on commitment contracts by applying it in a novel context (smartphone usage behaviors) and by investigating commitment contract effectiveness in relation to personality characteristics. This study further provides valuable insights for practitioners who try to help people decrease the time they spend on their smartphones.

Acknowledgements

I would like to thank my supervisor, Assistant Professor *Wiley Wakeman* of the Department of Management and Organization at the Stockholm School of Economics. Thank you for the outstanding support throughout the whole process of writing this thesis. From finding a suitable research topic, through conceptualization and execution of the experiment, to the very end of writing this paper, you always knew when to push me and when to slow me down. This thesis would not be the same without you!

Next, I would like to thank Professor *Magnus Johannesson* of the Department of Economics at the Stockholm School of Economics. Thank you for the many discussions about research designs, variable conceptualizations, and statistical analysis. You have been of enormous value and extraordinarily fast at giving me feedback when I needed it – even on red days!

I want to thank *Rebeka Karabeleska* for her continuous support – mentally and cognitively – throughout the last months. Thank you for everything!

Further, I would like to thank *Stina Hansson*, *Christoph Betting*, and *Felix Orth* for their support and feedback on my drafts. Especially Stina, thank you for the many hours you put into reading my texts and giving me feedback!

Lastly, to my niece *Nomi*, born on the very day that I defended this thesis. Thank you in advance for lighting up my life.

TABLE OF CONTENTS

1. INTRODUCTION	5
1.1 Problematization	8
1.2 Purpose of the study.....	10
1.3 Expected knowledge contribution.....	10
1.4 Delimitations	12
1.5 Research outline.....	13
2. THEORY	14
2.1 Literature review	14
2.2 Hypothesis generation.....	23
3. METHODOLOGY	30
3.1 Research approach	30
3.2 Preparation for the main study	32
3.3 Main study	34
3.4 Structure and data analysis	43
3.5 Quality considerations	45
4. ANALYSES AND RESULTS	48
4.1 Baseline characteristics.....	48
4.2 Hypothesis 1 – Testing commitment contract effectiveness in the context of smartphone usage.....	49
4.3 Hypothesis 2 – Relationship between individual’s future-self-continuity and commitment contract effectiveness	52
5. DISCUSSION	56
5.1 Effectiveness of commitment contracts in the context of smartphone usage.....	56
5.2 Relationship between individuals’ future-self-continuity and effectiveness of commitment contracts in the context of smartphone usage	58
5.3 Qualitative add-on: A quick look at good deeds performed.....	60
6. CONCLUSIONS	62
6.1 Theoretical contributions	63
6.2 Practical contributions	64
6.3 Limitations	65
6.4 Future research	70
7. REFERENCES	72
8. APPENDIX	81

LIST OF FIGURES

Figure 1 Stages of the experiment _____	35
Figure 2 Recruitment flow _____	42
Figure 3 x2 ANOVA Profile Plots; Commitment contract vs Goal Setting for High & Low FSC levels on Relative Phone Use Reduction _____	54
Figure 4 2x2 ANOVA Profile Plots; Commitment contract vs Goal Setting for High & Low FSC levels on Daily Goals Achieved _____	55

LIST OF TABLES

Table 1 Summary of hypotheses _____	29
Table 2 Baseline characteristics; Control vs Treatment on Full Sample, Finishers, Dropouts _____	48
Table 3 Commitment contract effectiveness vs simple goal setting _____	51
Table 4 Commitment contract effectiveness vs simple goal setting; assuming dropouts reverted to previous baseline usage _____	52
Table 5 2x2 ANOVA, descriptive statistics; Dependent variable: relative phone use reduction _____	53
Table 6 2x2 ANOVA, descriptive statistics; Dependent variable: daily screen time goals achieved _____	54
Table 7 Examples of good deeds performed _____	60

1. Introduction

A fast-paced, digitalized world enables large parts of the human population to live a life full of prosperity and abundance. Our lives are full of opportunities and there are so many different goals that we can set out for ourselves. I could be an actor. I could be a politician. I could be a musician, a tennis player, a teacher, or I could be a business student. Of course, achieving such kinds of larger goals usually requires longer series of steps to take. In that, meeting the small daily goals we set for ourselves seems to be important to achieve larger objectives that we pursue. In setting such daily goals, we make a lot of promises, not just to the people around us but mostly to ourselves: Tomorrow, I am going to get up at 6 am. I am going to exercise before I go to work. I am going to eat healthy. I am going to read more and spend less time on my smartphone. We all make big plans for our future selves. Unfortunately, many of us struggle to follow up on these intentions we have. In a TEDx Talk, behavioural economist Dan Ariely from the Duke University explained the underlying reasoning for our struggles of following up on our intended behaviours like this: “We all think that in the future, we are wonderful people. We will be patient, we will not procrastinate, we will exercise, we will eat well. The problem is that we never get to live in that future. We always live in the present” (TEDx Talks, 2011, 00:04:43). And in the present, our brains are very good at persuading us to why it is okay to snooze, to skip going to the gym, to eat that tasty Burger, or to pick up our smartphone just to scroll through social media or play mobile games. Sometimes our own irrationality pervades us from successfully following up on our intended behaviours.

This thesis does not serve to prophesize the dooming of our species due to human irrationality and our partial inability to achieve our goals. Often enough, we do achieve the goals we set ourselves. In doing so, we commonly draw on our self-control capacities (Fitzsimons & Finkel, 2010). Self-control can get us places. But exerting self-control can also be hard and we only have limited amounts of self-control (Gino, Schweitzer, Mead, & Ariely,

2011). This thesis attempts to shed light on how we might be able to help ourselves become better at achieving our goals. It might be possible to design our environment in a way that helps us follow up on our intended behaviours. To address such issues, researchers, practitioners, and individuals generally turn towards commitment devices. A commitment device is “an arrangement entered into by an individual with the aim of helping fulfill a plan for future behavior that would otherwise be difficult owing to intrapersonal conflict stemming from, for example, a lack of self-control” (Bryan, Karlan, & Nelson, 2010, p. 672). While there are countless forms of potential commitment devices, one subcategory of devices that has attracted increasing attention in research and in practice is the use of commitment contracts (Halpern, Asch, & Volpp, 2012). A commitment contract is a contract one makes with oneself or another party to achieve one’s future goals. It typically consists of two parts: 1) One commits to a certain behaviour in the future. 2) One commits to a cost in case one does not follow up on that planned behaviour. This mechanism is supposed to increase motivation to act upon one’s goals because one wants to avoid the agreed upon penalty (Halpern et al., 2012). In exploring potential ways to become better at achieving our goals, this thesis focuses on the use of commitment contracts.

However, not all people will respond identically to commitment contracts. Consequently, this thesis further examines how personality characteristics interact with commitment contracts in ways that increase or decrease their effectiveness. In that, this study takes a closer look at the interplay between individuals’ future-self-continuity and their ability to achieve goals with and without the use of commitment contracts. Future-self-continuity is a concept measuring how similar and connected we feel to our future selves. It has helped explain other individual behaviours and decision-making processes where self-control issues can arise. Individuals with a high future-self-continuity for instance seem to be much better at saving up for the future. (Hershfield, Garton, Ballard, Samanez-Larkin, & Knutson, 2009)

Self-control problems can overcome us in all kinds of ways – they can be related to our lifestyle, our health, private administrative matters, or our work and the projects we take on. This thesis focuses on the issue of smartphone usage. The pervasion of smartphones and their associated apps into almost every aspect of our lives is one of the greatest social trends of our generation (Kemp, 2018). The technological advances regarding mobile devices in the past years are remarkable and allow us to have all kinds of information, communication, or entertainment on the tip of our fingers. While the benefits mobile devices can bring are obvious for most, this development also brings potential downsides. Smartphones can be a great distraction in our everyday lives. We are always available, always online, and there is always *something* that draws on our attention (Anderson & Silver, 2019). But not only can smartphones distract us from potentially more important things or be an easy go-to option whenever we procrastinate; high smartphone usage also bears serious psychological health risks. Recent studies found that increased screen time can lead to loneliness, depression, and thoughts of suicide (Twenge, Joiner, Rogers, & Martin, 2017). Mobile devices and their applications are designed in a way to attract our attention and develop addictive behaviours (Andreassen, Pallesen, & Griffiths, 2017). The average adult's daily hours spent on mobile devices has more than doubled from 2012 (1.6 hours/day) to 2017 (3.3 hours/day) and continues to rise. Further drawing on a mobile consumer survey in the United States, many people think they spend too much time on their smartphones (39% of adults, 60% of 16 to 34-year-olds). 63 percent of respondents in that study indicated they had tried to limit their screen time, but only about half of those succeeded. Many are unable to decrease the time they spend on their phones and don't know how to achieve that goal. (Deloitte, 2017; Deloitte, 2018)

It is evident that the interrelation of smartphone usage and time-management is becoming more and more relevant. In this spirit, there may be a great deal of value to be gained from research on the effectiveness of commitment contracts regarding people's desire to decrease

their smartphone usage. The remainder of this introductory section serves to acquaint the reader with the addressed problem of this thesis, the purpose of the study, expected knowledge contributions, as well as distinctive delimitations.

1.1 Problematization

While there has been substantial interest in the research field of self-control and commitment devices for areas such as lifetime consumption and savings along with self-destructive health behaviours like obesity or smoking, there is surprisingly little research on commitment contracts to be found which are related to any other areas. Studies like Patterson's (2015), who investigated the effectiveness of commitment contracts regarding the use of distracting websites with student subjects in one of his classes, are rare. While the investigation of commitment contract effectiveness in the context of smartphone usage can thus be considered a relatively novel setting, further reasons for why this context was chosen to be investigated are outlined in the following.

Smartphone usage is a very up-to-date issue which is attracting rising interest and efforts from practitioners. Within the past years, more and more mobile applications have been developed that try to help individuals become more aware of their smartphone usage (Examples: Moment, Flipd). Many of these applications do so by tracking and providing information about individuals' screen time, i.e. the time a phone's display is turned on. Some go a step further and enable users to lock their phones or specific apps for some time. Others work with notifications and reminders or with goals to be set for a maximum time to spend on one's phone. Likewise, Apple introduced a new function called Screen Time in their iOS update from September 17th, 2018. Via this function, people can see their overall screen time as well as specific screen times for individual apps or different categories of apps such as entertainment, social networking, games, productivity, and more. Furthermore, Apple's Screen

Time function introduced the option to set time limits for single apps, app categories, or overall phone usage. It should be noted here, that once one reaches such a limit, one can simply click on “ignore for today”, and go on with one’s phone usage without any consequences. Most existing applications or functions simply inform people about their usage and, at the most, send a notification once one reaches a self-set limit. Reminders on these apps may work as an incentive to reduce screen time, but they are arguably weak ones with little effect. While there are many attempts to help people reduce their phone usage, there is not a lot of empirical work examining how effective these apps are in doing this. This thesis brings methodological rigor to the attempts of reducing screen time by drawing on commitment contracts. The conducted experiment intends to shed light on whether commitment contracts can be an effective tool in this context by not just monitoring screen time but incentivizing positive change in behaviour through the potential threat of a cost for not achieving one’s goal.

Looking at previous research on commitment contracts, one can also notice that examined research questions were mostly about the effectiveness of different types of commitment contracts. Some authors investigated deposit contracts, where participants put money on the line and only receive it back if they meet their goals (e.g. Volpp et al., 2008). Other scholars investigated positive incentives, where participants could gain money for achieving their goals (e.g. Charness & Gneezy, 2009). On another note, some studies focused on the designing of commitment contracts and on how likely individuals are to engage in them (Stedman-Falls, Dallery, & Salloum, 2018). However, previous research has failed to investigate whether commitment contracts, when in place, work better for some individuals than for others. Past trials have mostly been of clinical nature and considered different control variables such as age, gender, or income. Additional included variables were generally closely related to the context of the study, so that for instance studies on weight loss included weight and other health variables (Volpp et al., 2008; John et al., 2011). One may wonder about

potential insights to be gained from investigating whether, independent of the studied behaviour, the effectiveness of commitment contracts could generally be moderated by differences in individuals' psychological traits or concepts – for instance individuals' future-self-continuity.

1.2 Purpose of the study

This master thesis experiment has three main objectives. First, it shall examine the effectiveness of commitment contracts in the context of the modern-day issue of smartphone usage. Further, this study shall shed light on whether commitment contracts may be more effective for certain types of individuals than for others. Thus, the relationship between the concept of future-self-continuity and the effectiveness of commitment contracts will be investigated. While this individual measure may not be the only potential explanatory psychological variable with an impact on commitment contract effectiveness, there has been strong evidence for the connection between future-self-continuity and other self-control issues. Further elaborations on the hypothesized connection between commitment contract effectiveness and the concept of future-self-continuity are outlined in section 2 *Theory*.

Lastly, given the increasing public attention towards negative aspects of smartphone usage, this thesis aims to give important insights for practitioners (e.g. app developers) to develop and implement effective mechanisms which help individuals reduce the time they spend on their smartphones.

1.3 Expected knowledge contribution

As mentioned before, previous contributions to the field of self-control and commitment devices have mainly focused on the areas of lifetime consumption and savings as well as self-destructive behaviours. This study contributes to previous literature by taking the use of commitment contracts one step further and addressing the modern-day issue of smartphone

usage. The recent studies pointing towards a relationship between increased screen time and psychological health problems such as loneliness, depression, and thoughts of suicide (Twenge et al., 2017) support the importance of addressing smartphone usage behaviours with an intention to positively alter behaviour.

Furthermore, most studies on commitment contracts have focused on hard commitment contracts. This study adopts the use of soft commitment contracts where social efforts are demanded for failing to meet a goal. Generally, one can differentiate between two types of commitment contracts. While hard commitment contracts are of monetary nature such as penalties or rewards, soft commitment contracts are of a psychological or holistic nature (Bryan et al., 2010). In adopting a soft commitment contract, this thesis contributes to the scarce field of empirical studies about such soft commitments. An in-depth elaboration on why this study draws on soft and not hard commitment contracts follows in section *3.1.3 Research scope*.

Lastly, this thesis aims to enrich the research field by adding a potentially much-explaining layer of psychological traits to investigations on commitment contract effectiveness. In that, this thesis does not only explore whether commitment contracts help individuals achieve their goals, but also whether individuals may be better or worse at achieving their goals with and without commitment contracts depending on their future-self-continuity. Summarizing, this thesis shall serve the purpose of answering the following research question:

Can commitment contracts effectively help individuals in decreasing their smartphone usage, and does an individual's future-self-continuity effect these outcomes?

1.4 Delimitations

As this study is an exploration of the effectiveness of commitment contracts in the context of smartphone usage reduction, all findings shall only be taken as an investigation for this specific context. Of course, there are many other areas or contexts in which the effectiveness of commitment contracts could be investigated. Any results or insights from this study, however, should not be blindly transferred to other contexts, where one might wonder about the effectiveness of commitment contracts regarding study behaviour, work deadlines, diet enforcement, workout plans, or any other generally interesting area where self-control issues may play a role.

Also, all findings in this study are based on the use of soft commitment contracts. The results of this experiment will not allow to draw any conclusions about other forms of commitment contracts, for instance a deposit contract where individuals would put their own money on the line.

The investigation whether individuals' future-self-continuity may have an impact on individuals' ability to achieve their goals with and without commitment contracts, shall also not be generalized beyond this study. Any relationship to be found in this study should not be transferred to potential relationships between those two variables in other contexts than smartphone usage. Furthermore, this relationship is only investigated for the use of soft commitments.

One can think of many other psychological traits or concepts, for instance the big five personality traits (Barrick & Mount, 1991), that may help explain the effectiveness of commitment contracts. This experiment focuses on the concept of future-self-continuity and does not claim to investigate any other psychological variables that may be of interest.

1.5 Research outline

This thesis follows the subsequent structure of six main sections. This *Introduction* intended to familiarize the reader with important background information regarding the context and purpose of this study. Furthermore, it should have outlined this thesis's expected knowledge contributions as well as important delimitations.

The following *Theory* section acquaints the reader with existing literature that will help understand the nature and design of the performed experiment. It will conclude with a formulation of the hypotheses to be tested in this study.

Thereafter, the *Methodology* section outlines this study's methodological approach and argues for various methodological choices that had to be made in conducting this experiment.

Following that, the *Analyses and Results* section portrays the quantitative results from this experimental study.

The *Discussion* section will look at the implications of these results. This will be done by both relating the results to previous research and attempting to generate new insights in respect to previous gaps in research as outlined above.

In the *Conclusions* section, the key findings of this thesis as well as consequent theoretical and practical contributions are outlined. That will be followed by a critical discussion of this thesis' limitations and a suggestive outlook for future research.

2. Theory

Issues of self-control, goal achievement, and time-management are certainly relevant for any individual or organization. Consequently, motivation and goal setting are core components of management research. It can be argued that this thesis falls into the field of behavioural economics, and the following section shall give an important overview of relevant concepts and previous studies that help put the performed experiment into perspective.

2.1 Literature review

The art of possessing self-control seems desirable for most people around the world. Previous research has shown that self-control allows one to save more money for retirement (Thaler & Benartzi, 2004), quit smoking (Volpp et al., 2009; Halpern et al., 2015), stay healthier by keeping track of weight (Volpp et al., 2008; John et al., 2011), or perform well in school by studying for upcoming tests (Levitt, List, Neckermann, & Sadoff, 2016; Patterson, 2015). Many positive things come out of self-control and often people wish they had more of it to stick to their plans and achieve their goals. The desire to improve one's self-control and address this lack of discipline is an underlying reason to why this area has been studied by psychologists and behavioural economists. Luckily, many studies have attempted finding ways to deal with shortfalls of self-control. One of these investigated ways is the use of commitment devices.

Commitments or commitment devices can take many different forms. The idea is to take a binding action of some form ahead of time to influence future actions and choices (Bryan et al., 2010). Examples of this would be signing up for a program to lose weight or to study. Strotz (1955) was the first to link this type of behaviour to economics. He formalized a theory of commitment to show that commitment devices can be important determinants of economic outcomes. Strotz (1955) laid the foundation for what was later labelled as hyperbolic

discounting – a concept that will be further explained in *2.1.1 Hyperbolic discounting*. He showed that when people's discount functions are non-exponential, meaning that they care about their future selves and future payoffs and don't discount them in a way that follows an exponentially increasing slope, they prefer to constrain their own future choices. He also found that individuals are often willing to engage in quite costly commitments.

However, research has acknowledged that an individual's decisions are often suboptimal and recognises errors and judgmental biases that contribute to such suboptimal behaviours. According to Thorgeirsson & Kawachi (2013), previous research in this field has mainly focused on retirement savings and consumption among individuals while more recent studies try to address health problems such as smoking and obesity. This literature review addresses all the above-mentioned areas within the field to familiarize the reader with the current state of related works.

2.1.1 Hyperbolic discounting

Economists tend to deal with the impatience of individuals by assuming an exponential discounting of utility over time. In other words, people seem to prefer rewards sooner rather than later and incur costs as late as possible. While there is truth and evidence for this view, it has been developed to reflect a state of mind which differs somewhat from traditional exponential discounting. Hyperbolic discounting or quasi-hyperbolic discounting is a theory and model developed by Laibson (1997) that addresses this shortfall. It is considered to be a cornerstone for present research in the field of behavioural economics.

Hyperbolic discounting is time-inconsistent and suggests that the valuation of a payoff or future event falls swiftly during the first couple of periods, such as the first couple of days, to then flatten out and fall much slower for later periods. According to Laibson (1997, p. 443) "hyperbolic discount functions induce dynamically inconsistent preferences, implying a motive for consumers to constrain their own future choices". Laibson (1997) considers the individual

as a sequence of “temporal selves” making choices for the future with the assumption that they can foresee their future self-control problems. He argues that people are in demand of commitment devices to limit the choices of their future selves and shows for instance that consumers sometimes prefer illiquid assets to constrain their future selves.

O'Donoghue & Rabin (1999) also modelled the individual future valuations as time-inconsistent. They study self-control problems and label the phenomenon as present biased preferences. Similar to Laibson (1997), they divide people into two groups regarding their view on their own future self-control problems and call them sophisticated or naive. They consider activities which involve immediate rewards and immediate costs and find that naive people procrastinate immediate cost activities and preproperate (i.e. to do too soon) immediate-reward activities. They also find that sophisticated individuals do not procrastinate costs to the same extent but somewhat more surprisingly they tend to have stronger preproperation of rewards. It is argued that present biased preferences only slightly hurt the naive individuals due to postponing costs while it can severely hurt the sophisticated due to heavily prioritising immediate rewards. How this distinction between naive and sophisticated individuals may be relevant for this study's investigations will be further discussed in *2.2 Hypothesis Generation*. For now, one can note that the evidence for a time-inconsistent discounting by individuals is widely supported by researchers. Furthermore, the notion that people make sup-optimal decisions over time is rooted in previous research on the topic of self-control.

2.1.2 Commitment contracts

In more recent studies in the field of behavioural economics, commitment contracts as a form of commitment device have extensively been used to impose self-control. Commitment contracts have taken many forms in several different contexts, of which the most notable ones are outlined in this section. Throughout many studies, commitment contracts have proven to considerably help in achieving desired behaviour in the long-term, including productivity and

motivation at work (e.g. Kaur, Kremer, & Mullainathan, 2015; Cui, Shi, & Xu, 2017), saving for the future such as pensions (Ashraf, Karlan, & Yin, 2006; Thaler & Benartzi, 2004), and health related applications like weight loss (Volpp et al., 2008) or smoking cessation (Giné, Goldberg, & Yang, 2010). Furthermore, Bryan et al. (2010) studied how different types of commitment devices compare in terms of effectiveness. They also established the distinction between hard and soft commitments to investigate differences between respective effectiveness for commitments of monetary and psychological nature. Among their suggestions was that soft commitments, in contrast to hard commitments, seem to be particularly good for designing effective commitment contracts.

Even though previous studies in the field of mobile screen time are scarce, one study that is somewhat related to this thesis was done by Patterson (2015) on how to improve outcomes for students in an online course. The study shows that commitment contracts seem to work better than previously tried devices. The students initially committed to time limits on distracting websites, reminders when they engage in distracting activities online and a way of blocking certain websites when it was time to study. The students in the treatment group spent more time working on the course, received higher grades, and were far more likely to complete the course. It was concluded that commitment tools that reduce procrastination have a significant impact on student performance. This study also suggested that students in the treatment groups found unproductive time much less enjoyable (Patterson, 2015). While some may see that as proof for that commitment contracts come at a cost, one can also argue that the commitment devices worked as intended by making the time spent on unproductive websites less pleasant.

2.1.2.2 Health - Exercise & Smoking

Another key area for the application of commitment contracts within behavioural economics is the altering of self-destructive behaviours to improve health. One of the popular

applications is to reduce obesity as has been studied by Volpp et al. (2008) and John et al. (2011). These studies contribute by underlining the possibility for people to commit ahead of time to achieve long-term interests. In John et al. (2011), subjects in the control group entered the program with monthly weigh-ins only. The treatment groups additionally received one of two financial incentive plans to lose weight. The financial incentives in this program involved deposit contracts, which are commitment contracts where subjects can lose their own money if they do not lose weight. The study showed significant weight loss for the treatment groups. Unfortunately, subjects regained weight after the study was concluded, suggesting that commitment contracts may only have long term effects for as long as they are active. (John et al., 2011)

Halpern et al. (2015) used a similar method with financial punishments. They state that commitment contracts where money is deposited and only returned upon success have substantial appeal as a method of altering destructive behaviour within health. Another example of this is the study by Della Vigna & Malmendier (2006) in which they examined how subjects acted when they committed to paying in case they decided to skip going to the gym. Their results state that some subjects are overconfident and overestimate attendance along with the fact that the rational expectation hypothesis is insufficient at properly explaining consumer preferences. This insight regarding individual's partial inaccuracy of predicting future self-control problems will become relevant again when the hypotheses regarding the interplay of commitment contract effectiveness and individuals' future-self-continuity are discussed.

Similar types of financial commitment contracts and incentives have proven to work in smoking cessation studies. One such study was conducted by Volpp et al. (2009). The experiment was conducted in a large company and the treatment group had significantly higher rates of smoking cessation, enrolment, and completion of the program.

Commitment contracts of different forms are shown to be very effective as seems to be evident from many recent studies. It has also been stated that both direct financial incentives and deposit contracts seem to be more acceptable among subjects than previously believed (Stedman-Falls et al., 2018).

2.1.2.3 Education

As touched upon previously, another popular application of commitment contracts is within the field of education. Levitt et al. (2016) performed a series of field experiments including many thousands of subjects from primary and secondary school. They found that one can significantly influence performance within education, both through financial and non-financial incentives. However, non-financial incentives seem to decline in power for older students, suggesting that the effectiveness of financial incentives increases with age relative to non-financial. From this insight, one may be able to make some assumptions about the expected effectiveness of soft commitments (non-financial) used in this study, where the sample consists of mostly adults.

Most interestingly, Levitt et al. (2016) also find that rewards for students are consistent with hyperbolic discounting, namely that rewards handed out with a delay are weak incentives and do not significantly increase investment into studies. As grades and traditional rewards for education come with a delay, the current educational system may not be optimal for maximising individual performance. Furthermore, this stresses the importance for commitment contracts to be designed in a way so that rewards or costs incur as soon as possible after a goal is met or failed. For this master thesis experiment, the cost of performing a good deed was expected to be carried out on the very next day, i.e. the soonest possible. Contrary to previous studies, Levitt et al. (2016) do not see any gains from framing rewards as costs or losses.

While all these studies outlined above broadly looked at the effectiveness of commitment contracts, they did not look at individuals' personalities. Personality affects many performance

indicators and organizational outcomes (Rothmann & Coetzer, 2003) as well as individuals' discount rates (Bartels & Rips, 2010). Therefore, considering personality differences in studying commitment contracts could be important.

2.1.3 Future-self-continuity

The concept of future-self-continuity (FSC) is a fundamental building block for this thesis. It is based on some people finding it more difficult to postpone payoffs than others. Hershfield et al. (2009) wrote a paper on the subject testing the FSC hypothesis, namely that individual's level of FSC, i.e. their similarity and connectedness to their future-selves, is related to saving rates. High levels of FSC mean that one possesses a lower discount rate for future payoffs and a more consistent perception of one's present- and future-self. The study validated the FSC index and found evidence that higher levels of future-self-continuity correlate with higher savings.

In general terms, it was found that “specifically, when the future-self shares similarities with the present self, when it is viewed in vivid and realistic terms, and when it is seen in a positive light, people are more willing to make choices today that may benefit them at some point in the years to come” (Hershfield, 2011, p.30).

Further, Blouin-Houdon & Pychyl (2015, p.51) found that “individual differences in perceived similarity to one's future-self predicts procrastination such that participants who experienced higher future-self-continuity in ten years [...] reported fewer procrastination behaviours”. These findings serve as a strong foundation for the expectation of individuals with a high future-self-continuity being better at exerting self-control and achieving their smartphone usage reduction goals than individuals with a low future-self-continuity. It shall soon be discussed why these findings do not provide clear expectations regarding the effectiveness of commitment contracts for High and Low FSC individuals, though. Consequent studies have further shown that the concept of future-self-continuity can be related to other

issues of self-control and even that individuals' FSC levels can be manipulated so that subsequent behaviour shows more future-oriented considerations (van Gelder, Hershfield, & Nordgren, 2013; van Gelder, Luciano, Weulen Kranenberg, & Hershfield, 2015). This thesis investigates whether the concept of future-self-continuity also helps explain issues of self-control in the context of goal achievement and commitment contracts.

Studies on concepts closely related to the concept of future-self-continuity provide further supporting evidence for a connection between individuals' psychological connection to their future selves and intertemporal choices (Bartels & Rips, 2010) and that an individual's awareness of self-control problems can help mitigate problems of the traditional present-biased preferences towards harmful addictions and overconsumption (O'Donoghue & Rabin, 2002). The introduction of individuals' future-self-continuity in a study on commitment contract effectiveness intends to address the lack of personality investigations in previous commitment contract research. The concept of future-self continuity has been proven to explain how good people are at exerting self-control in other areas and thus may be of explanatory value in the context of smartphone usage and commitment contract effectiveness.

2.1.4 Potential downsides of commitment contracts

While the various ways in which commitment contracts can help mitigate issues of self-control have been outlined previously, there are however certain downsides to commitment contracts and optimising for the future. Particularly, one may mention that these mechanisms can be costly. Charness & Gneezy (2009) also suggest that certain incentives can be counterproductive, for instance in fields such as education. They can result in the individual focusing too much on certain scores or the measured metrics rather than the more holistic approach of truly doing something for one's best. Furthermore, many studies have discovered that the behaviour of an individual is not fundamentally changed by commitment contracts or incentives and often reverts when they are removed. Some studies do however still support the

fact that these tools can be used to overcome the initial resistance to performing an action (Angrist & Lavy, 2009; Angrist, Lang, & Oreopoulos, 2009; Kremer, Miguel, & Thornton, 2009).

Additionally, Falk & Kosfeld (2006) created a principal-agent game where the principal imposes certain minimum performance requirements on the agent. They found that the productivity and performance of agents is in fact reduced in response to the imposed control due to signals of distrust and limitation of choice. Consequently, it is important that people opt-in to commitment contracts by their own choice. Involuntary participation could cause serious problems because some individuals may see the intervention as a sign of mistrust or paternalism.

2.1.5 Concluding remarks

While the studies and papers outlined above conclude that commitment contracts in fact help individuals to alter their behaviours, Rogers, Milkman, and Volpp (2014) also suggest that expanding commitment contracts beyond just monetary losses will lead to an increase in popularity, such as committing to something with friends and relatives, perhaps through utilising social media. Having failure lead to disappointment among peers is a very cost-efficient device. Looking for more than just financial incentives and leveraging newer technology will hopefully help increase engagement in the real world. (Rogers et al., 2014)

In conclusion, researchers seem to be unanimous that commitment contracts possess great conceptual strength and that they can be used to significantly improve savings or health, reduce smoking, and can contribute in reaching goals set in many other areas. One new area will be investigated in this thesis. Generally, commitment contracts have shown substantial promise throughout previous experiments.

2.2 Hypothesis generation

Building on previous research, this section gives a brief outline of the expected subject behaviour. Arguments for tested hypotheses in this experimental study will be presented and the predictions of each hypothesis are expressed. Explanations for expected behaviours will draw on previous literature.

2.2.1 Suggested effectiveness of commitment contracts in the context of smartphone usage

The first area of discussion is linked to the effectiveness of commitment contracts in helping people reduce the time they spend on their phones and achieve their daily screen time goals. Previous studies on commitment contracts which specifically investigated the effectiveness of commitment contracts compared to simple goal setting or other incentive systems, looked at varying but distinctive measures in evaluating the effectiveness of commitment contracts. Studies on weight loss measured weight loss over certain periods of time, such as 16 weeks (Volpp et al., 2008) or 32 weeks (John et al., 2011). These weight-loss studies set monthly goals to be achieved, based on which subjects could lose money if not achieved. Furthermore, these studies also analysed effectiveness in terms of total average weight loss per incentive group at the end of the study. Studies on smoking cessation measured smoking cessation over time periods of six, nine, or twelve months (Volpp et al., 2009; Halpern et al., 2015). Studies on exercising behaviour measured the amount of times subjects went to the gym (Charness & Gneezy, 2009). Patterson (2015) measured the time students spent working on the course, their grades, and completion rates of the course. Since there is no historical evidence of any experiment on smartphone usage reduction, this thesis cannot build on previous literature in choosing a measure of effectiveness, but steps up suggesting two measures to be used in this experiment and in potential future research. First, this thesis measures the effectiveness of commitment contracts in the number of days for which subjects

achieve their goal, i.e. stay below their new daily screen time goal. This can be comparable to monthly weigh-ins regarding weight loss studies such as in Volpp et al. (2008). Furthermore, this thesis looks at relative reduction in smartphone screen time, i.e. a measure in percentage terms, which is calculated based on a subject's average screen time during the week prior to starting the experiment. This measure may be comparable to total average weight loss per incentive group like in John et al. (2011) or to students' time spent working on a course like in Patterson (2015). Reasons for why a relative instead of an absolute measure was applied in this study will be further outlined in *3.3.3 Measures and conceptualizations*.

It can be argued that excessive smartphone usage is comparable to other unhealthy behaviours, such as smoking or obesity. Therefore, one may argue to expect subject behaviour to be comparable to that of previous studies looking at commitment contracts in relation to those behaviours. All studies previously mentioned in this paragraph show significant results suggesting that commitment contracts are more effective than simple goal setting, regardless the context of investigation. Consequently, it can be expected that commitment contracts will likely have a greater impact on screen time reduction and achieving daily screen time goals than simple goal setting.

One can theorize that commitment contracts will help motivate the subjects to cut their smartphone usage to a greater degree than simple goal setting. This may occur through several mechanisms. In simple terms, commitment contracts are expected to help individuals with this self-control problem. They do this through inducing extrinsic motivation, when intrinsic motivation is low or ineffective. While this is the simple premise of commitment contracts, there are additional motivating factors worthy of note.

The inclusion of a soft incentive could help increase motivation by simply providing an explicit cost to smartphone usage past a certain point. This could help improve subjects'

motivation as they would be able to put an expected effort value to the cost of using their phone. Consequently, the total cost of smartphone usage should be higher for the treatment group.

Secondly, commitment contracts could help increase motivation by bringing costs of smartphone usage into the present. The implicit costs of excessive smartphone usage can come in a range of forms. Whether it is through reducing the amount of time one has finishing a piece of work or the cost of foregoing healthier behaviours such as exercising, these costs are predominantly in the future. Following from Strotz' (1955) hyperbolic discounting research, individuals are likely to have difficulties in constraining current behaviours for benefits in the distant future. A commitment contract brings the explicit costs of smartphone usage into the present – at least at the margin of the daily limit. This would be in line with Levitt et al.'s (2016) results, suggesting that punishments or rewards given without a delay work as stronger incentives and have a larger impact on subjects than punishments or rewards that come with a delay.

Concluding, previous research provides large theoretical support for the hypothesizing that commitment contracts can positively impact individuals' attempts in reducing their smartphone screen time and achieving their daily screen time goals. Consequently, this study will investigate the following hypothesis:

H1: Compared to subjects exposed to simple goal setting, subjects exposed to commitment contracts will be significantly better at reducing their smartphone usage (H1a) and achieving their daily screen time goals (H1b).

2.2.2 Suggested relationship between individuals' future-self-continuity and effectiveness of commitment contracts in the context of smartphone usage

In analysing the relationship between individuals' future-self-continuity and the effectiveness of commitment contracts, this section will look at four different forms of interplays separately. It shall discuss 1) the relationship between individuals' FSC levels and exposure to goal setting, 2) the relationship between individuals' FSC levels and exposure to commitment contracts, 3) potential differences between Low FSC individuals exposed to goal setting and Low FSC individuals exposed to commitment contracts, and 4) potential differences between High FSC individuals exposed to goal setting and High FSC individuals exposed to commitment contracts. Thereafter, an all-explaining hypothesis for the four relationships is formulated.

Regarding the expected relationship between individuals' FSC levels and exposure to simple goal setting, previous research clearly suggests that individuals with higher future-self-continuity are better at exerting self-control and delaying rewards (Hershfield et al., 2009). One may portray the use of one's smartphone as a reward, since it can still our hunger for entertainment, communication, or distraction. Furthermore, Blouin-Hudon & Pychyl (2015) discovered that low levels of future-self-continuity are related to high levels of procrastination. Smartphone usage, especially in the form of social media and entertainment could be considered as a form of procrastination. Consequently, it can be expected that, when exposed to simple goal setting, individuals with higher future-self-continuity will be more successful in decreasing their smartphone usage and achieving their daily screen time goals than individuals with lower future-self-continuity.

Regarding the expected relationship between individuals' FSC levels and exposure to commitment contracts, previous literature does not allow us to create strong and specific expectations for what to observe. On the one hand, it could be argued again that individuals

who feel more similar and connected to their future selves may be better at exerting self-control, i.e. delaying rewards (Hershfield et al., 2009) and not procrastinating (Blouin-Hudon & Pychyl, 2015). On the other hand, commitment contracts may just be the type of mechanism that helps individuals with low future-self-continuity to overcome these self-control issues. Due to the lack of research connecting the effectiveness of commitment contracts to any psychological concepts, one cannot presume to know whether individuals with high FSC levels still outperform individuals with low FSC levels. Potentially, the mechanism of a commitment contract bridges the lack of self-control that individuals with low future-self-continuity generally have by introducing a cost to the potential non-achievement of a goal. Hershfield et al. (2009) suggest that greater future-self-continuity means an individual fears future losses stronger. This notion may further be supported by Bartels & Rips (2010), arguing that individuals prefer costs to occur after changes in psychological connectedness to one's future-self. Thus, the introduction of an almost-immediate cost might have more impact on individuals with lower FSC levels because the threatening cost tackles the present self at all times, i.e. every day while exposed to commitment contracts. Consequently, it can be expected that, when exposed to commitment contracts, subjects with higher future-self-continuity will not be differently successful in decreasing their smartphone usage and achieving their daily screen time goals from individuals with lower future-self-continuity.

Regarding potential differences between Low FSC individuals exposed to goal setting and Low FSC individuals exposed to commitment contracts, any differences between the two can be attributed to the commitment contract. Given the previous proposition that Low FSC individuals exposed to goal setting may struggle in exerting self-control, potential differences arise depending on whether the outlined mechanisms introduced through the commitment contracts help those individuals exposed to them bridge their lack of self-control. If that proves to be the case, one can expect that Low FSC individuals are more successful in decreasing their

smartphone usage and achieving their daily screen time goals when exposed to commitment contracts compared to simple goal setting.

Lastly, potential differences between High FSC individuals exposed to goal setting and High FSC individuals exposed to commitment contracts may or may not be expected. Even though previous literature generally suggests commitment contracts to be more effective than simple goal setting, one may encounter insignificant differences for High FSC individuals. This belief is based in the argument that a high level of FSC means subjects have a lower discount factor for future payoffs (Hershfield et al., 2009). Thus, the benefit from bringing costs of smartphone usage into the present may not have as large an impact as it does on subjects with low future-self-continuity. Furthermore, we already established the expectation that simple goal setting may be more effective for subjects with high FSC than for subjects with low FSC because they are able to see the longer-term benefits of these goals, and able to exert greater self-control to achieve these goals. Therefore, the comparative benefits of commitment contracts would be lower. Alongside this, a potential that subjects with a high FSC already have healthier or lower levels of smartphone usage could be considered. Thus, imposing a 30% reduction may be a far harder goal for them to achieve as the marginal gain of each minute of smartphone usage could be higher for these subjects. This could decrease the perceived effectiveness of commitment contracts. If their smartphone usage is already low, the marginal utility of use could be above the penalty imposed by the commitment contract. A short analysis on the differences in subjects' average screen time prior to the experiment was conducted to see if there are any a-priori differences between High FSC individuals and Low FSC individuals. No significant differences could be found. Concluding, it can be expected that individuals with high future-self-continuity exposed to commitment contracts are not more successful in decreasing their smartphone usage and achieving their daily screen time goals than individuals with high future-self-continuity exposed to simple goal setting.

In combining these four expected relationships between exposure to commitment contract or simple goal setting and high or low future-self-continuity, an overarching hypothesis can be generated which suggests that individuals' FSC levels may act as a moderator in predicting commitment contract effectiveness. A negative general relationship between individuals' FSC levels and commitment contract effectiveness is expected to be found. Consequently, this study will investigate the following hypothesis of a moderating effect:

H2: Commitment contracts have a stronger positive impact on relative smartphone usage reduction (H2a) and daily screen time goals achieved (H2b) for individuals with low future-self-continuity than for individuals with high future-self-continuity.

Table 1 Summary of hypotheses

H1: Compared to subjects exposed to simple goal setting, subjects exposed to commitment contracts will be significantly better at reducing their smartphone usage (H1a) and achieving their daily screen time goals (H1b).
H2: Commitment contracts have a stronger positive impact on relative smartphone usage reduction (H2a) and daily screen time goals achieved (H2b) for individuals with low future-self-continuity than for individuals with high future-self-continuity.

3. Methodology

In the following section the methodology will be presented, starting with an outline of the research approach. Thereafter, methodological decisions and designs for pre-studies as well as the main study will be presented. This section will then finish with an elaboration of the data structure and analysis, lastly discussing quality considerations.

3.1 Research approach

This section about the research approach will be further broken down into research reasoning, research design, and research scope.

3.1.1 Research reasoning

In this thesis, exploring whether commitment contracts help individuals achieve their goals in the context of smartphone usage reduction and whether the helpfulness of such commitment contracts depends on individuals' levels of future-self-continuity, a deductive research methodology is used. Even though an abductive or inductive research method could bring new insights to the field, a deductive approach is suitable to address the limited time in conducting the study, and to avoid risk (Dudovskiy, 2016). Moreover, a deductive approach is the most common approach in related research and allows to investigate causal relationships. Existing theories were used to design the research study and consequent hypotheses which are tested empirically. (Bryman & Bell, 2015)

3.1.2 Research design

To answer the research question “*Can commitment contracts effectively help individuals in decreasing their smartphone usage, and does an individual's future-self-continuity effect these*”

outcomes?”, a quantitative method is chosen, using a longitudinal experimental design. In that, this study can be categorized as a framed field experiment with field context in commodity and task (Harrison & List, 2004). The quantitative method, using data to examine theory, is in alignment with a deductive approach (Dudovskiy, 2016; Saunders, Lewis, & Thornhill, 2012). Moreover, it allows for understanding of the phenomena in question through quantifiable answers from close-ended questions in a questionnaire, number’s driven statistical results, and objective data presentation (Creswell, 2014).

The experimental design included one treatment and one control group. Participants were randomly allocated into the two groups (Mertler, 2016). The experimental design is beneficial as it allows the studying of things that participants might not be aware of, such as underlying psychological processes, which is highly relevant for this study (Schwarzer & Fuchs, 1996). The design also benefits from that it may allow for (to some degree) generalizable and transferable results to the population represented by the sample, and because it allows for treatment manipulation (Bryman & Bell, 2015). Furthermore, also in line with deductive approaches, the design enables to investigate correlations and causal effects (Perdue & Summers, 1986). In testing the hypotheses, principles of an intention-to-treat analysis are applied, meaning that non-compliance with the assigned treatment does not exclude participants from the sample and subsequent analyses (Alshurafa et al., 2012). By allowing for an unbiased comparison between treatment and control group, an intention-to-treat analysis represents a more conservative basis for the measure of effect sizes (Hardin, 1998; Heritier, Gebiski, & Keech, 2003).

3.1.3 Research scope

This quantitative experimental study focuses on the effectiveness of commitment contracts in the context of smartphone usage. This choice of context was based on *interest*, in

the field of commitment contracts as well as in individual behaviors regarding smartphone usage, *relevance*, in the increasing amount of everyday issues related to phone usage as well as first studies pointing towards negative health effects from excessive screen time (Twenge et al., 2017), and *practicability*, relating to timely and monetary constraints of this study. Moreover, this thesis only studies the use of soft commitment contracts. By applying soft instead of hard commitment contracts, this thesis contributes to the scarcity of empirical studies on soft commitment contracts. In addition, soft commitment contracts are chosen due to their nature, where participants put in social efforts in the case of failing to meet a goal, instead of monetary punishments. Monetary punishments, such as in the form of deposit contracts, would have increased difficulties in recruiting, thus limiting the sample size.

3.2 Preparation for the main study

In preparing for the main study, a few uncertainties regarding the experimental design had to be addressed. These uncertainties centered around the study population. Consequently, two pre-studies were performed to address two separate issues regarding the study population.

3.2.1 Pre-study 1

It became clear early in the process that participation in the experiment had to be restricted to iPhone users. This decision was made because the Operating System of any iPhone updated after September 17th, 2018, has a pre-installed function called Screen Time, which allows users to observe the exact time they spent on their phone during any day of the past seven days. That function was invaluable for the conduction of this experiment, and thus, only participants with an iPhone running on an operating system of iOS 12 or higher could participate. The information from the Screen Time function should serve as a benchmark for how much time participants spent on their phone on average (=baseline usage). The information should further be used to calculate a new daily screen time goal for any participant

throughout the experimental week, and it should help individuals check up on their performance throughout the experimental week. Users of smartphones with other operating systems such as Android, would have had to download an app that does that service for them. Even if they did that, no historical phone usage benchmarks would have been available for the study unless those individuals tried to use their phones normally for one week to record baseline usage, leading to heterogeneous difficulties between iPhone and Android users that would be difficult to adjust for. To avoid such difficulties, a pre-study was conducted to investigate whether there are any differences in demographics or other relevant concepts between iPhone users and users of smartphones running on other operating systems.

For this pre-study, 60 participants were recruited through public posts on Facebook and LinkedIn. 32 individuals reported to use iPhones and 28 reported to use Androids. To test for differences in baseline characteristics between iPhone users and Android users, independent samples t-tests were used for continuous variables. Pearson's χ^2 test or Fisher's exact test was used for categorical variables, as appropriate. The results showed no significant differences between the two groups in any of the baseline characteristics, including demographics and future-self-continuity levels. Consequently, the results of this pre-study suggested no differences between iPhone users and Android users. Thus, it could be concluded that the main analysis may focus on iPhone users without limiting generalizability of the results to the general population. The results of this analysis can be found in appendix 1.

3.2.1 Pre-study 2

Due to conflicting pressures regarding sample size and response quality, as well as limitations in time and monetary resources, a decision had to be made on how to recruit participants. The main decision criteria were the response and data quality. Various options were considered, of which two ended up being pursued. For a short test period, participants were recruited both personally by the researcher and via a service-provider. The personal

recruitment was based on social media posts by the researcher as well as approaching people on the streets. The second option made use of Prolific Academics, an online-platform where individuals from all over the world can take surveys for monetary rewards. For both types of recruitment, around 45 people were recruited during this test period. All individuals were exposed to the main study experiment consisting of two surveys to start and end the experimental week. It was decided that the data collection for the main study should solely focus on personal recruitment, i.e. leveraging social media and recruiting people on the street. While the participants from Prolific Academics were more likely to complete the experiment (i.e. return for survey 2) than participants recruited personally, the data quality in both surveys (survey 1 and survey 2) showed problems for the Prolific Academics responses. A large proportion of the participants failed to pass attention checks or included data that was inconsistent. For instance, 22 out of 40 respondents in this recruitment group failed to correctly re-report their new daily screen time goal given to them in survey 1, when asked for it in survey 2. Even though some of the responses would have been useful, it was decided that the general data quality in this recruitment group is not sufficient to use it as a main recruitment technique. The recorded responses via social media & street recruitment were further included in the main analysis, as this type of recruitment remained as the main recruitment approach.

3.3 Main study

For the main study, an experimental between-subjects research design was used. To test the hypotheses of this thesis, an experiment was conducted in a one-week session for each participant. A participant's session was split into seven rounds, each round lasting for a 24-hour calendar day. To view and assess subjects' natural phone usage, they were asked to check the Screen Time function within the settings of their iPhones. The experiment was conducted on volunteers who, given their completion of the entire experiment, automatically entered a lottery draw in which one out of every 15 participants was randomly selected to receive 200

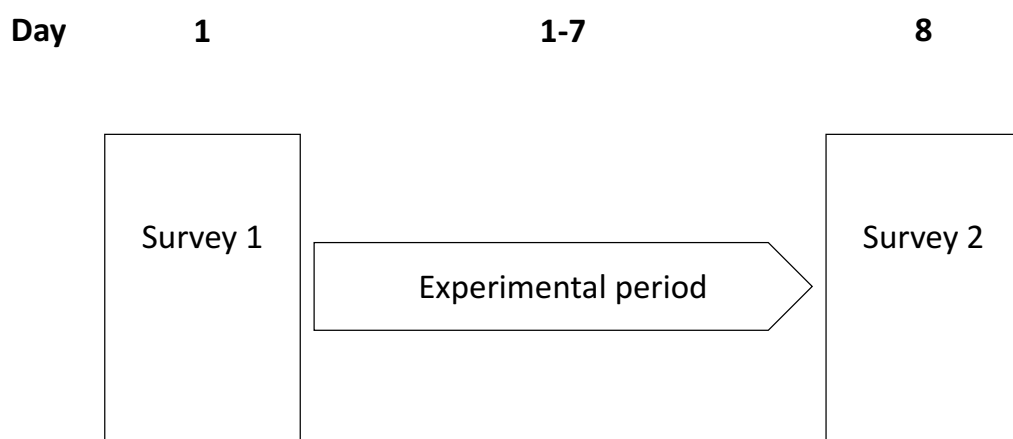
SEK. Participants were asked to accept the terms of this experiment, whereby they will submit self-reported phone usage data.

Both treatment and control group participants underwent the same general stages of the experiment. The experiment started with survey 1, where participants answered questions regarding their future-self continuity, demographics, and their baseline phone usage. Further, they were given a new daily screen time goal.

After completion of survey 1, the experimental week started immediately. For this time, participants were asked to try to decrease their smartphone usage. In that, all participants were given a new daily screen time goal, which they should try not to exceed.

After completion of the experimental period, participants were asked to answer a second survey, in which daily phone usage times from the experimental week were reported. Additional follow-up questions were recorded in survey 2. The entire process for all individuals is represented in figure 1. The following sections outline specific design choices made for this study.

Figure 1 Stages of the experiment



3.3.1 Experimental groups and treatment design

Most of the instructions given to participants were identical for both control and treatment group. In survey 1, subjects allocated to both groups were asked to try to reduce their smartphone usage for the experimental period. They were given a new daily screen time goal, which they should try not to surpass during the following seven days. These new goals were individually calculated by multiplying individuals' average smartphone usage of the seven days prior to submitting survey 1 with a factor of 0.7. Thus, the new daily screen time goal for all participants is equal to a 30% reduction compared to their baseline smartphone usage. Each day, they should check if they had succeeded on staying below their goal the day before. All participants were further asked to use an iPhone function called *App Limits* that provides a dismissible notification five minutes before as well as exactly when they reach their screen time limit for the day. The only differing instructions for control and treatment group were regarding the goal setting for the experimental week.

Goal setting in the control group meant that participants were simply asked to set themselves the goal of decreasing their smartphone usage and to try to stay below their new daily screen time goal throughout the next week. Subjects in the control group were further instructed that there are no formal penalties or rewards for achieving or failing their daily screen time goals.

Goal setting in the treatment group involved a short explanation about the nature and functional principle of commitment contracts. Participants were then asked to sign such a commitment contract stating that they will aim to achieve their goal. For that, an electronic signature function in Qualtrics was used. The commitment contract specified that for every day they should fail to stay below their limit, they will perform a good deed to someone else on the following day and write down the good deed to report it together with their actual phone usage in survey 2.

3.3.2 Survey design

Prior to starting the main study, understandability of both surveys in general as well as each question in specific were pilot-tested by a small number of respondents (n=6). This quality check is in accordance with suggestions given by Bryman & Bell (2015). All questions were perceived as clear, so that no adjustments had to be made.

3.3.2.1 Survey 1

The first survey (see appendix 2) consisted of five main blocks, of which only the last block was subject to randomization to allocate participants into treatment and control group. All blocks consisted of explanatory texts and structured questions with closed answers only.

The first block introduced participants to the experiment. In a descriptive manner, this block presented the outline of the experiment and acquainted participants with the purpose of research, the time required for participating, as well as the associated risks and benefits of participation.

In the second block, participants were first informed about the concept of future-self-continuity and then asked four structured questions, which constitute the future-self-continuity measure.

The third block was composed of various questions regarding demographic data. Respondents were also asked to insert an e-mail address so that they could be contacted for the second survey. Lastly, this block featured one attention check question.

Block number four asked participants to report the average daily time they spent on their phone during the previous seven days. Based on this average, which participants were asked to report based on the *Screen Time* function in their iPhones, a new daily screen time goal was computed. Participants were then asked to try to decrease their smartphone usage and stay below that new daily goal for the next seven days. Participants were further instructed to

activate the function *App Limits* in their iPhones and to write down their screen time goal as well as all actual screen times for each day during the experimental week.

For the last block, Qualtrics randomly assigned participants to either a control or a treatment condition. In the control condition, participants were simply asked to set themselves the goal of staying below their daily screen time limit for the next seven days. In the treatment condition, participants were informed that one aspect of this study is to investigate the helpfulness of commitment contracts in smartphone usage. The general mechanism of a commitment contract was described. Participants were then asked to sign such a commitment contract stating that they intend to try to stay below their daily screen time limit for the next seven days. The commitment contract further included that for every day that they fail to do so, they will do something good for someone else on the following day, and that they will note down their good deeds to report them in survey 2. For both conditions – control and treatment – the survey ended with a short summary regarding the next steps of this experiment. Section 3.3.3 *Measures and conceptualizations* will further describe all items used to measure variables in detail.

3.3.2.2 Survey 2

The second survey (see appendix 3) consisted of two blocks. Only the last block differed slightly between the treatment and control group. Participants were sent a treatment or control version of the second survey based on which condition they were randomly assigned to in survey 1. Generally, all blocks consisted of explanatory texts and structured questions with closed answers.

The first block served to welcome back participants to the experiment and to remind them about what to expect for the second survey. Participants were further reminded that only upon completion of this survey will they automatically enter a lottery to win 200 SEK.

In the second block, participants were asked to re-report the daily screen time goal they had during the experimental week. Lastly, they reported their actual screen times for every single day of the experiments before leaving their email address again, so that responses from survey 1 could be matched with responses from survey 2. For treatment group participants, this last block additionally consisted of an open question asking individuals to report the good deeds they had performed for the days they failed to stay below their goal.

3.3.3 Measures and conceptualizations

As mentioned before, this section describes all items used to measure variables in detail.

3.3.3.1 Future-self-continuity

The construct of future-self-continuity was measured in accordance with Hershfield et al. (2009) who established this four-item index to measure individuals' connectedness to their future selves. Participants were asked to imagine themselves ten years from now and answer four related questions. The first two questions asked participants to choose the circle pair that best describes how similar and how connected they feel to their future-self ten years from now. Seven graphics were displayed for both questions, which showed circle pairs overlapping to different degrees, comparable to a seven-point Likert scale. One circle represents one's current-self, the other represents one's future-self. Overlaps ranged from no overlap to almost complete overlap. The last two questions then asked how much participants like and care about their future-self ten years from now. Both questions were rated on a seven-point Likert scale ranging from "*don't like at all*" / "*don't care at all*" to "*completely like*" / "*care a great deal*". All four items were combined into an index as suggested by Hershfield et al. (2009) and adopted by further studies (e.g. Hershfield, 2011; Blouin-Hudon & Pychyl, 2015; van Gelder et al., 2013, van Gelder et al. 2015). Cronbach's α for the four items was 0.58. While this lies below the general acceptable level, previous research that has used the concept of future-self-continuity, such as Hershfield et al. (2009), show similar levels ($\alpha=0.57$). Thus, the index was used for

further analyses and not adapted. To assess the impact of future-self-continuity on the effectiveness of commitment contracts, participants were then categorized as individuals with high and low FSC levels, to see if the effectiveness of commitment contracts varies with participants' FSC levels. For this split, the median FSC level was used as a reference point to divide subjects into subgroups. While there are discussions about whether median or mean splits should be used for such approaches (Allen, 2017), in the case of this study both would have resulted in the same division of participants.

3.3.3.2 Baseline smartphone usage

Participants' baseline smartphone usage was a self-reported measure. It was measured in minutes. Participants were instructed to access historical data in their iPhone settings which provides accurate details about their daily screen time during the past seven days.

3.3.3.3 New daily screen time goal

To provide specific goals throughout the experimental week, participants were given a new daily screen time goal. New goals were calculated by multiplying participants' previous average screen time (baseline usage) with a factor of 0.7. This measure was given in minutes and rounded up when not resulting in a full number. For instance, if a participant reported a baseline smartphone usage of 137 minutes per day, her new daily screen time goal for the experimental week was set at 96 minutes per day.

3.3.3.4 Outcome variable: Relative smartphone usage reduction

One of the two outcome variables in this study was relative smartphone usage reduction. Because of the absence of previous self-control literature in the context of smartphone usage, this new outcome variable had to be developed. It was derived and adapted from Volpp et al. (2008) and John et al. (2011), who measured commitment contract effectiveness in the context of weight loss via total weight loss in kilograms after 16 and 32 weeks, respectively. While absolute smartphone usage reduction in minutes may appear the most similar to the measure of

absolute weight loss in kilograms, it was decided to look at relative smartphone usage reduction in percentages. In comparison to weight loss, screen time changes can happen much faster and much more drastically. Furthermore, in the context of screen time, the baseline usage has a much stronger impact on absolute reductions than the baseline weight in a weight loss study. Consequently, a relative instead of an absolute measure of screen time reduction was used.

3.3.3.5 Outcome variable: Daily screen time goals achieved

The second outcome variable used was the number of daily goals participants achieved throughout the experimental week. Again, this measure draws upon Volpp et al. (2008) and John et al. (2011), who both use monthly weigh-ins for which weight-loss goals were given to participants. These weigh-ins serve both as a binary outcome variable (goal achieved vs goal not achieved) and as an anchor towards the commitment contract itself. Individuals would thus lose money if they failed to achieve a monthly weigh-in goal. In the context of smartphone usage, using daily screen time goals seemed an appropriate counterpart to monthly weigh-ins. Since smartphone usage behaviors can fluctuate much faster than individuals' weight, daily goals are applied in this study. Daily goals further allow to establish an integer-variable ranging from zero to seven, which allows for more variation in this outcome variable than in binary outcome variables used in other commitment contract studies on behaviors such as smoking cessation (see Halpern et al. 2015; Volpp et al., 2009).

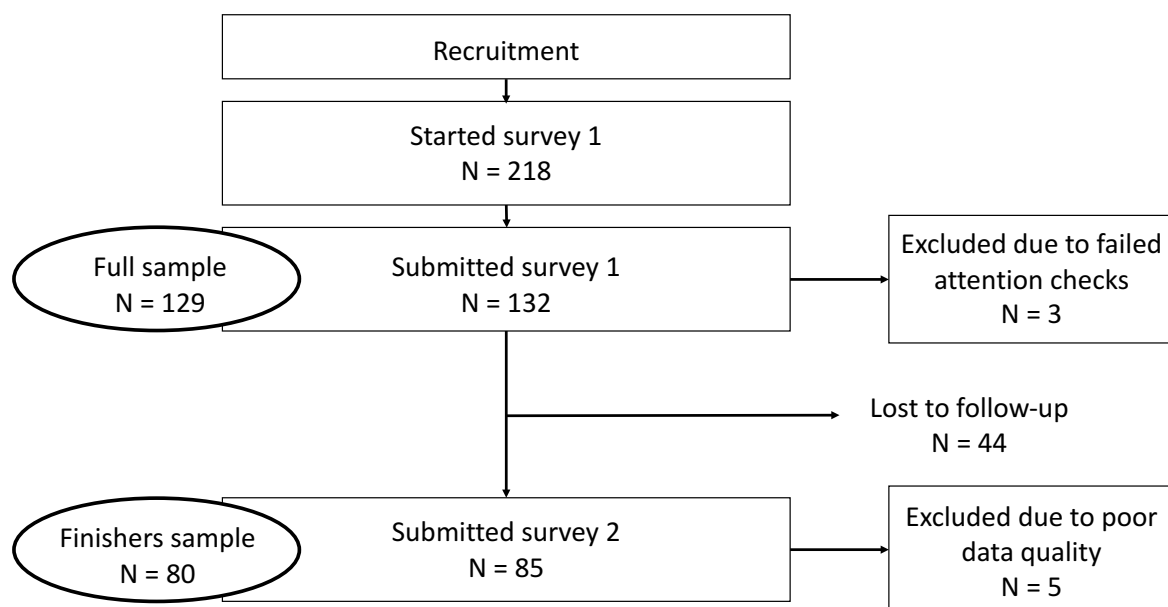
3.3.3.6 Demographics

To control for potential differences between groups, the following demographic variables were measured (provided answer options in the survey are indicated in brackets): gender (male, female, other), age (in years), marital status (single, in a relationship, married, separated, widowed/divorced, don't want to disclose), employment status (employed, self-employed, unemployed, student, other), working hours per week (measured in 10 hour spans from 20 to 60 hours; options to answer below 20 hours or above 60 hours).

3.3.4 Sampling approach and data collection

Participants were recruited during a five-week time window. They could start the experiment at any time, i.e. when they were recruited. The study sample of this experiment consists of volunteers who were recruited via the following methods. 72 volunteers were recruited via online recruitment. For that, public posts were made on Facebook and LinkedIn. Furthermore, 60 volunteers were recruited on the street. For that, the author approached individuals at public places in Stockholm such as Odenplan, T-Centralen, Stockholm School of Economics, or Stockholm University. Altogether, 218 individuals started survey 1. Out of those, 132 individuals finished the first survey and started the experiment. 85 participants finished the experiment by submitting survey 2 after their experimental week had ended. Three participants who started the experiment but did not finish it had to be excluded because of failed attention checks in survey 1. Furthermore, five participants who finished the experiment had to be excluded because of poor data quality or failed attention checks in survey 2. Consequently, this study could build its analyses on a full sample size of 129 subjects and a finisher-sub-sample size of 80 subjects. Figure 2 provides an overview of the recruitment flow.

Figure 2 Recruitment flow



All participants who completed the experiment (i.e. submitted survey 2) entered a lottery in which one out of every fifteen participants was randomly selected to win 200 SEK.

To collect all data for survey 1 and 2, Qualtrics was used. Thus, all responses were online-based and self-completed by participants. For the first survey, participants accessed the questionnaire via a link or via scanning a QR code. For the second survey, emails were sent out to every participant on exactly the day on which their individual experimental period had just ended. For the second survey, up to four reminder emails were sent out to participants. The use of reminder emails is in accordance with a suggestion for how to deal with missing outcome data in an intention-to-treat analysis by White, Horton, and Carpenter (2011).

3.4 Structure and data analysis

All data for both surveys was recorded in Qualtrics. For survey 1, two identical surveys were used to identify how participants were recruited (online / street). For survey 2, two separate surveys were used so that participants in control and treatment group could be sent the appropriate version of survey 2. All data sets were exported from Qualtrics and opened in Microsoft Excel, where all responses could be sorted by e-mail address so that survey 1 and survey 2 data could be combined for all participants. This dataset was directly transferred to IBM SPSS 25 to be used for consequent analyses.

3.4.1 Data checks

After excluding invalid responses based on survey 1, the full sample consisted of 129 participants. A post-hoc analysis was conducted to ensure that there were no statistically significant differences between the different recruitment populations. Invalid responses based on survey 2 were excluded from the finishers-sample but remained in the full sample because their reported baseline characteristics were acceptable. Their inappropriate data from survey 2

was deleted, so that they were treated as dropouts. Independent sample t-tests, Pearson's χ^2 test, and Fisher's exact test were used to test for differences in baseline characteristics. Independent sample t-tests and a 2x2 ANOVA were used for hypotheses testing. Prior to all tests, the data of relevant variables was checked for normal distribution via the Kolmogorov-Smirnov test, suggesting no significant deviations from normal distribution for any of the tested variables. Further, no problems with multicollinearity could be detected. Levene's test for homogeneity of variances showed significance for the outcome variable of daily goals achieved ($p=0.009$). Thus, independent sample t-tests not assuming equal variances were used for tests with this variable. No outliers were excluded from the data set. For all analyses, a confidence interval of 95% was used to establish significant findings.

3.4.2 Recoding of variables

For the future-self-continuity variable, the four-item scale was combined into an index variable. The working hours per week variable was recoded so that each data point, which referred to a span such as 21-30 hours, was transformed into the middle value of that span (e.g. 25.5). The commitment-contract-signature-ID was recoded into a binary variable, referring to data in the treatment group with 1, and the control group with 0. For the 2x2 ANOVA, a binary variable was computed, which used the future-self-continuity variable as an input to put out 1 for High FSC individuals and 0 for Low FSC individuals. A median split was conducted here (min.: 2.25, max: 7, median: 5.25). The outcome variable daily screen time goals achieved was created by counting for how many of the seven days, the reported screen time was equal to or below a subject's screen time goal. The outcome variable relative smartphone usage reduction was created by first taking the average of the seven reported daily screen times and then calculating the following equation to result at:

$$\text{relative smartphone usage reduction (in \%)} = 1 - (\text{new average} / \text{baseline average}).$$

3.5 Quality considerations

The following sections address relevant aspects of data quality in this study. As for all quantitative studies, the methods of data collection and the collected data itself should be discussed in relation to reliability and validity implications (Bryman & Bell, 2015).

3.5.1 Reliability

From the conceptual start of this experiment to the data analysis and assessment of results, this study was designed and executed with an aim to hold up to high standards of reliability and transparency. In that, all steps of the experiment were critically planned beforehand and meticulously aligned with relevant concepts included in the theoretical framework as well as with indispensable data collection standards. It was tried to minimize any potential impact on participant behavior that could come from interaction with the researcher during recruitment or the follow-up e-mails that were sent out for survey 2. These arrangements were made to increase reliability of the experiment and allow for repeating or replication of the study (Bryman & Bell, 2015). This thesis further reports all methodological decisions and steps as detailed as possible to provide full transparency and allow others to make reliability conclusions for themselves.

Since all data was collected via questionnaires, reliability of all constructs used in this experiment depends on the accuracy and correctness of participants' responses. Regarding the baseline phone usage variable and the relative smartphone usage reduction variable, reliability can be considered high. This is due to the technical nature of how these data points are observed. Accurate and objective screen time measures were recorded via the *Screen Time* function in participants' iPhones. However, a chance for human error in the process of reading those data points and transferring them into the Qualtrics surveys cannot be ruled out. Consequently, the general reliability of reported screen times could be decreased due to

mistakes in reporting or calculating the hours-and-minutes measure from the phone into a minutes-only measure in Qualtrics. Lastly, since screen times are self-reported, the data quality also hinges upon participants' honesty in reported screen times. The same reasoning applies for the outcome variable daily screen time goals achieved, since this measure was entirely dependent on the reported screen times.

3.5.2 Validity

In discussing this study's validity considerations, the following sections differentiate between measurement validity, internal validity, and external validity (Saunders et al., 2012).

3.5.2.1 Measurement validity

In preparation of the experiment, assuring high measurement validity was a main objective. Consequently, this study attempted to make use of well-established and previously used scales and measurement practices for all concepts where this was possible (Saunders et al., 2012). Due to the absence of any established outcome variables in the context of smartphone usage reduction, two new measurement scales were developed. The development of these scales was based on similar outcome measures in previous commitment contract studies. As mentioned above, reliability of these measures can be considered high. Due to their technical and objective origin, it can further be argued that daily screen time goals and relative phone usage reduction capture their relative concepts as intended. However, alternative ways of measuring effectiveness in smartphone usage reduction could be considered. Therefore, future research should investigate different measurements and their respective validities. Only then will it be possible to make positive statements about the validity and consequent appropriateness of the newly developed measures in this study, such as a potential superiority of a relative over an absolute measure for smartphone usage reduction.

3.5.2.2 Internal validity

Due to randomization of participants into treatment and control group, any differences in baseline characteristics between the two groups are caused by chance. Even though smartphone usage behavior can be impacted by external factors, randomization also allows to assume that both groups were exposed to similar likelihoods of being affected by such factors. However, the limited sample size in this experiment harms internal validity to some degree, so that the possibility of random outcomes cannot be ruled out with the highest level of confidence. Overall, internal validity in this study is sufficient so that differences in outcome variables can largely be attributed to the treatment, i.e. exposure to a commitment contract.

3.5.2.3 External validity

The conducted experiment can be classified as a framed field experiment. Field context in commodity and task increase generalizability of results compared to a conventional lab experiment or an artefactual field experiment (Harrison & List, 2004). Consequently, findings can be generalized in that the observed behaviors are likely to happen in a non-experimental real-life context as well. Subjects were asked to report data on their actual smartphone usage. Other than having set screen time goals for one week, participants went on with their lives as usual. While this aspect increases external validity of this study, it must also be mentioned that the sample cannot be considered representative of the overall population. Certain limitations persevere in drawing statistical inferences about the general population based on the given sample characteristics. Consequently, external validity is high in relation to the context of the studied behavior, but limited in relation to the population the study sample aims to represent. (Saunders et al., 2012)

4. Analyses and results

This section starts with a presentation of participants' baseline characteristics including a short analysis of the differences between treatment and control group in baseline characteristics for the full sample, the finishers-sample (i.e. participants who finished the experiment), and the dropouts-sample (i.e. participants who were lost to follow up). Thereafter, conducted analyses and respective results are outlined in the order of the hypotheses.

4.1 Baseline characteristics

To test for differences in baseline characteristics between groups, independent samples t-tests were used for continuous variables. Pearson's χ^2 test or Fisher's exact test was used for categorical variables, as appropriate.

Table 2 Baseline characteristics; Control vs Treatment on Full Sample, Finishers, Dropouts

	Full Sample		Finishers		Dropouts	
	Control n=67	Treatment n=62	Control n=42	Treatment n=38	Control n=25	Treatment n=24
Gender						
Male	31 (45.6)	27 (42.8)	22 (52.4)	17 (44.7)	9 (36.0)	10 (41.7)
Female	37 (54.4)	34 (54.8)	20 (47.6)	20 (52.6)	16 (64.0)	14 (58.3)
Other	0	1 (1.6)	0	1 (2.6)	0	0
Age (sd)	25.78 (7.77)	25.03 (5.37)	25.71 (7.83)	24.95 (5.14)	25.88 (7.82)	25.16 (5.84)
Marital status						
Single	30 (44.1)	26 (41.9)	21 (50.0)	17 (44.7)	9 (36.0)	9 (37.5)
Relationship	27 (39.7)	33 (53.2)	15 (35.7)	20 (52.6)	12 (48.0)	13 (54.2)
Married	9 (13.2)	1 (1.6)	4 (9.5)	1 (2.6)	4 (16.0)	0
Divorced/widowed	2 (2.9)	1 (1.6)	2 (4.8)	0	0	1 (4.2)
Don't want to disclose	0	1 (1.6)	0	0	0	1 (4.2)
Employment status						
Student	46 (67.6)	45 (72.6)	30 (71.4)	30 (78.9)	16 (64.0)	15 (62.5)
Employed	18 (26.5)	14 (22.6)	9 (21.4)	6 (15.8)	8 (32.0)	8 (33.3)
Self-employed	3 (4.4)	1 (1.6)	2 (4.8)	0	1 (4.0)	1 (4.2)
Unemployed	0	2 (3.2)	0	2 (5.3)	0	0
Other	1 (1.5)	0	1 (2.4)	0	0	0
Working hours (sd)	30.4 (16.0)	30.1 (18.3)	32.6 (15.8)	32.3 (18.2)	27.4 (16.3)	27.3 (18.4)
FSC (sd)	4.87 (0.9)	5.04 (0.86)	5.12 (0.93)	4.97 (0.88)	4.83 (0.88)	4.77 (0.81)
Baseline usage (in minutes)	178.5 (91.9)	191.9 (92.5)	169.2 (77.8)	184.6 (80.2)	194.2 (111.7)	203.5 (119.2)

Note: numbers in brackets indicate percentages within the given group unless standard deviation is indicated by (sd)

As participants were allocated into treatment and control group randomly, by definition, any potential differences between the two groups are caused by randomness. Randomization ensured comparability between both groups for further analyses where the full sample is used.

For the finishers-sub-sample, participants in the treatment and the control group showed no statistically significant differences in any of the baseline characteristics. This suggests comparability between both groups for further analyses where the finishers-sample is used.

For the dropouts-sub-sample, participants in the treatment and the control group also showed no statistically significant differences in any of the baseline characteristics. This provides support for that individuals who dropped out from the experiment did not do so based on which group they were randomly assigned to. Furthermore, Pearson's χ^2 test confirms that dropout rates for both groups are similar (treatment = 39%, control = 37%; $\chi^2: p=0.870$).

In comparing the finishers-sample with the dropouts-sample, potential differences between dropouts and finishers can be analysed. Participants who dropped out and participants who finished the experiment showed no statistically significant differences in baseline characteristics. This implies that no selection effects regarding who dropped out or completed the experiment can be found.

4.2 Hypothesis 1 – Testing commitment contract effectiveness in the context of smartphone usage

Hypotheses 1a) and 1b) suggest that subjects exposed to commitment contracts are better at reducing their smartphone usage (H1a) and achieving their daily screen time goals (H1b) than subjects exposed to simple goal setting. In the main analysis, only the 80 participants who completed the entire experiment by submitting survey 2 are included. This is due to the large proportion of dropouts in this study. Roughly 63% of participants who started the experiment by submitting survey 1 (129 subjects) finished the experiment and submitted survey 2 (80 subjects) after the experimental week. This study intended to adopt strict intention-to-treat

principles. In that, problems arose because of the large proportion of missing outcome data. Even for smaller proportions of missing outcome data, scholars admit that there is no common practice to deal with such a problem (Alshurafa et al., 2012). However, this study attempts to address this by conducting a main analysis on the finishers sample followed by a sensitivity analysis on the full sample in which assumptions about dropouts are made. This is in line with how White et al. (2011) suggest dealing with missing outcome data in an intention-to-treat analysis. However, it is worth mentioning, that any kind of assumptions for missing outcome data can be considered as strong and artificial, which is why the main analysis in this thesis is based on the finishers-sample for which full datasets are available. Furthermore, within the finishers-sample, this analysis uses intention-to-treat principles by not excluding non-compliers (individuals who did not perform any good deeds even though they failed a daily screen time goal at least once). Consequently, this analysis attempts to follow intention-to-treat principles by performing a sensitivity analysis on the full sample and by including all participants who finished the study no matter whether they complied to the instructions given in the treatment group.

For the main analysis of hypothesis one, two independent samples t-tests were conducted to compare the two outcome variables for treatment and control group. Because Levene's test for homogeneity of variances showed significance for the outcome variable of daily goals achieved ($p=0.009$, relative smartphone usage reduction: $p=0.345$), independent samples t-tests not assuming equal variances were conducted. The results presented in table 3 suggest that the treatment group did have higher means for both outcome variables relative phone usage reduction (treatment=19.32%, control=11.46%, $t=-1.565$, $p=0.122$) and daily goals achieved (treatment=3.22, control=2.19, $t=-2.158$, $p=0.035$). However, only the difference in daily screen time goals achieved is significant at the 5% confidence level. Hence, hypothesis 1 is partially supported, finding no significance for 1a) but significance for 1b). We find effect sizes

according to Cohen’s *d* of 0.32 for relative phone usage reduction and of 0.48 for daily goals achieved.

Table 3 Commitment contract effectiveness vs simple goal setting

Dependent variable	Control n=42	Treatment n=38	<i>t</i>	<i>p</i>
Relative phone usage reduction in % (Standard deviation)	11.46% (20.55%)	19.32% (24.05%)	-1.565	0.122
Daily screen goals achieved – x out of 7 (Standard deviation)	2.19 (1.73)	3.22 (2.39)	-2.158	0.035*

Note: *significant at the 5% significance level

H1: Compared to subjects exposed to simple goal setting, subjects exposed to commitment contracts will be significantly better at reducing their smartphone usage (H1a) and achieving their daily screen time goals (H1b).

Partially supported

To provide a more conservative estimate of the differences between the groups, a sensitivity analysis on the previous results was performed using the full sample, where dropouts were treated as having reverted to their previous baseline usage as reported in survey 1. As mentioned before, this approach is in line with how White et al. (2011) suggest dealing with missing outcome variables in an intention-to-treat analysis. Further, this approach is also similar to how John et al. (2011) dealt with individuals lost to follow-up in their commitment contract weight-loss study, where they treated dropouts as having reverted to baseline weight. For this sensitivity analysis, the assumption of dropouts having reverted to baseline usage implies that all dropouts’ outcome variables for relative reduction and daily goals achieved consist of 0% and 0 out of 7 days, respectively.

Again, two independent samples t-tests were conducted to test the hypotheses, comparing the two outcome variables for treatment and control group. Because Levene’s test for homogeneity of variances suggested heteroscedasticity-tendencies for both outcome variables

(relative smartphone use reduction: $p=0.061$; daily goals achieved: $p=0.001$), independent samples t-tests not assuming equal variances were conducted. The results for this analysis are presented in table 4. One can find non-significant differences between treatment and control group for relative phone usage reduction (treatment=11.84%, control=7.18%, $t=-1.375$, $p=0.168$) as well as for daily goals achieved (treatment=1.95, control=1.37, $t=-1.532$, $p=0.128$). The corresponding effect sizes, according to Cohen's d , are 0.24 and 0.28, respectively. Given the assumption and its implication for that roughly 37% of our outcome data equals zero in this case, the mean treatment effects decrease in this sensitivity analysis.

Table 4 Commitment contract effectiveness vs simple goal setting; assuming dropouts reverted to previous baseline usage

Dependent variable	Control n=67	Treatment n=62	t	p
Relative phone usage reduction in % (Standard deviation)	7.18% (17.13%)	11.84% (20.99%)	-1.375	0.168
Daily screen goals achieved – x out of 7 (Standard deviation)	1.37 (1.73)	1.95 (2.43)	-1.532	0.128

Note: *significant at the 5% significance level

4.3 Hypothesis 2 – Relationship between individual's future-self-continuity and commitment contract effectiveness

Hypothesis two intended to investigate the relationship between individuals' future-self-continuity and commitment contract effectiveness. In that, it was hypothesized that the concept of future-self-continuity helps understand for which individuals a commitment contract may have a stronger impact, i.e. is of more help in reducing smartphone usage and achieving daily screen time goals. It was suggested that commitment contracts have a stronger positive impact on individuals with low future-self-continuity than on individuals with high future-self-continuity. Hypothesis two was tested using a 2x2 ANOVA test. In a 2x2 ANOVA analysis, such a moderating effect is indicated by the interaction term between FSC level and commitment contract dummy variable. For the following analyses, the finishers-sample was

used. This decision was based on the same reasoning as given for why this sub-sample was used for the main analysis in hypothesis one.

The analysis tests whether, generally, commitment contracts have a stronger positive impact on individuals with low FSC than on individuals with high FSC. The descriptive statistics of this analysis for both outcome variables can be found in table 5 and 6. Furthermore, the profile plots for both outcome variables resulting from the 2x2 ANOVA are presented in figures 3 and 4. The results suggest a tendency towards the predicted relationship between FSC levels and commitment contract effectiveness for both variables, however, the interaction terms between FSC levels and commitment contract dummy variable in both analyses for relative phone use reduction ($p=0.418$) and daily goals achieved ($p=0.675$) are non-significant. Hence, empirical evidence does not support the suggested general relationship between individuals' FSC levels and commitment contract effectiveness.

Table 5 2x2 ANOVA, descriptive statistics; Dependent variable: relative phone use reduction

FSC	Treatment	Mean	sd	n
Low	Goal setting	4.23%	22.08%	18
	Commitment contract	16.04%	28.68%	20
	Total	10.45%	26.11%	38
High	Goal setting	16.77%	17.31%	24
	Commitment contract	20.31%	21.50%	18
	Total	18.29%	19.05%	42
Total	Goal setting	11.39%	20.24%	42
	Commitment contract	18.07%	25.29%	38
	Total	14.56%	22.86%	80

Note: *sd* = standard deviation;

Table 6 2x2 ANOVA, descriptive statistics; Dependent variable: daily screen time goals achieved

FSC	Treatment	Mean	sd	n
Low	Goal setting	1.83	1.465	18
	Commitment contract	3.10	2.292	20
	Total	2.50	2.024	38
High	Goal setting	2.46	1.888	24
	Commitment contract	3.33	2.497	18
	Total	2.83	2.186	42
Total	Goal setting	2.19	1.728	42
	Commitment contract	3.21	2.361	38
	Total	2.68	2.103	80

Note: *sd* = standard deviation;

Figure 3 x2 ANOVA Profile Plots; Commitment contract vs Goal Setting for High & Low FSC levels on Relative Phone Use Reduction

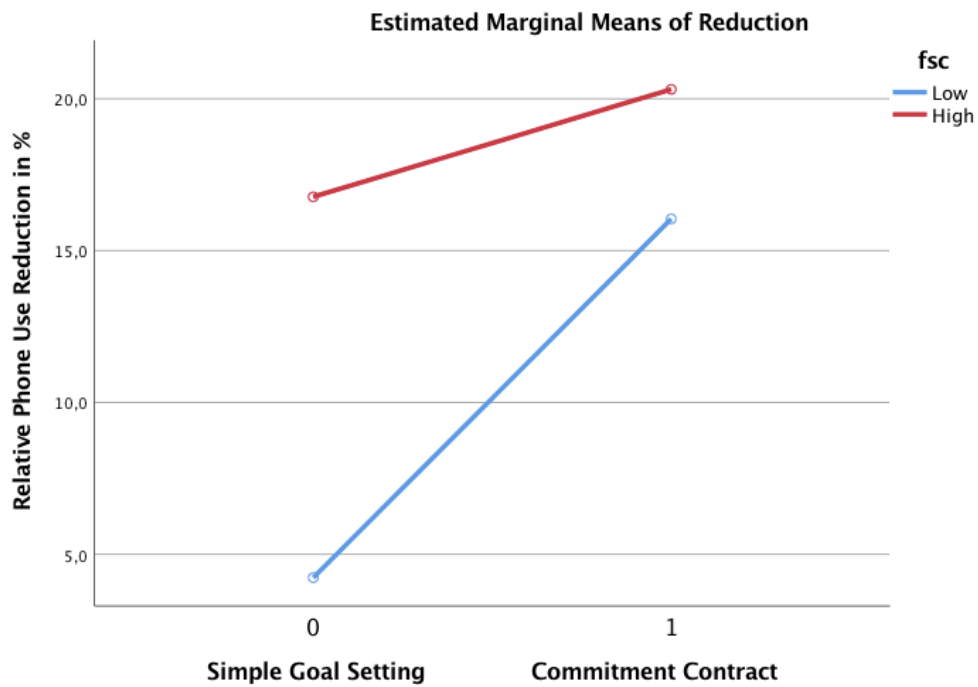
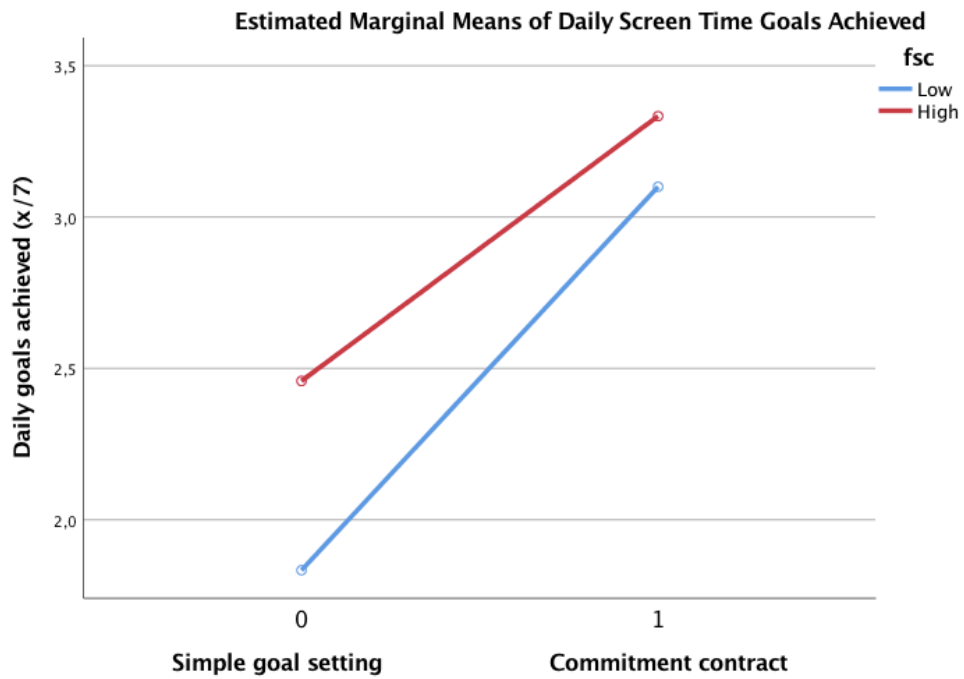


Figure 4 2x2 ANOVA Profile Plots; Commitment contract vs Goal Setting for High & Low FSC levels on Daily Goals Achieved



H2: Commitment contracts have a stronger positive impact on relative smartphone usage reduction (H2a) and daily screen time goals achieved (H2b) for individuals with low future-self-continuity than for individuals with high future-self-continuity.

Not supported

5. Discussion

This section provides an overview of the results presented above. In that, the findings regarding the two hypotheses will be discussed separately. The following discussions intend to specify what can be learned from this experiment and attempt to connect the findings to previous literature and theory from the fields of commitment contracts. Furthermore, the results will also be discussed regarding what cannot be learned from this experiment. It will be drawn upon potential explanations for why the research design and most of all the sample size resulted in problems to make high-power statistical analyses, suggesting that also non-significant findings may be considered potential contributions to the field.

5.1 Effectiveness of commitment contracts in the context of smartphone usage

The first aim of this experiment was to investigate the effectiveness of commitment contracts in the context of smartphone usage. It was hypothesized that commitment contracts help individuals spend less time on their smartphones. Consequently, the relative effectiveness of commitment contracts compared to simple goal setting was tested. Effectiveness was measured in the two outcome variables relative phone usage reduction and daily screen time goals achieved. The results indicate that, compared to when people simply set the goal to decrease the time they spend on their phones, commitment contracts could not significantly help individuals in their relative smartphone usage reduction. However, the results suggest that commitment contracts do serve as an effective tool to help individuals in staying below (i.e. achieving) daily screen time goals. Thus, hypothesis one was partially supported, because only support for sub-hypothesis 1b) was found. These findings are consequently partially in line with previous literature regarding commitment contract experiments, where the superior effectiveness of commitment contracts has been proven in various contexts.

To explain why this experiment may only have been able to partially support the hypothesis about commitment contract effectiveness in the context of smartphone usage, one could start by looking at the logical interplay between the two measured outcome variables. One can easily argue for a connection between the two. If a person successfully stays below her daily screen time goal more often than another person, that person is also likely to show a larger relative reduction in smartphone usage. Nonetheless, commitment contracts only had a significant effect on the daily goals achieved in this experiment. It could be speculated for why that might be. Potentially, participants exposed to commitment contracts are better at staying below their daily screen time goals but then show stronger fall-back behaviours, meaning that once a daily limit is surpassed, they spend much more time on their phones for the rest of that day because there is no additional cost or punishment. For this study, in contrast to for instance weight loss studies, an individual's daily performance is not dependent on the daily performance from the day, week, or month before. Since the goals were anchored on a daily basis, once such a goal was surpassed, from the participant's perspective it may not make a difference anymore whether the limit is surpassed by a few or many minutes. However, with the data and insights from this study, it is not possible or appropriate to make any positive statements about such psychological explanations. Any serious attempts at explaining how such detailed behaviours may come together should be subject to future research that is specifically directed at exploring more detailed processes.

Furthermore, looking at the effect sizes and significance levels for both outcome variables, the differences are not too extensive. Section 6 *Conclusions* will touch upon the limitations of this study which may have made it difficult to find significant results for hypothesis one in general and thus stress the significance of finding significant results for at least one of the two outcome variables. Future research could attempt to re-investigate the use

of commitment contracts in this context with variations in the research design and larger sample sizes.

By looking at the behaviour of smartphone usage, this experiment expands the commitment contract literature regarding applicable contexts and at least partially supports the potential effectiveness of commitment contracts in helping individuals to spend less time on their phones.

This study is further partially in line with previous literature suggesting that soft commitment contracts can be effective tools in changing behaviour (Bryan et al., 2010). Since the field of soft commitment contracts is relatively scarce, these results help broadening this field regarding behavioural context and design of soft commitment contracts. Due to the design of the commitment contract and the associated cost for failing to achieve a daily goal, this thesis showed that even soft commitment contracts where the cost is not enforced by default but hinges upon the compliance of the individual, can be effective tools in shaping behaviours.

5.2 Relationship between individuals' future-self-continuity and effectiveness of commitment contracts in the context of smartphone usage

The second focus of this study was to investigate whether the psychological construct of future-self-continuity could help explain which individuals benefit most from the use of commitment contracts. In that, the relationship between individuals' future-self-continuity and commitment contract effectiveness was explored. It was hypothesized that commitment contracts have a stronger positive impact on individuals with low future-self-continuity than on individuals with high future-self-continuity. The 2x2 ANOVA results provide no support for this relationship. Due to the absence of any previous research connecting commitment contract effectiveness with personality characteristics, this study serves as a first expedition into such waters. The analysis and results on this relationship cannot provide support for a

moderating effect of future-self-continuity on commitment contract effectiveness. Thus, individuals' future-self-continuity did not explain who might benefit most from the use of commitment contracts. Furthermore, these non-significant results do not allow for suggesting that previously established relationships between the concept of future-self-continuity and other self-control issues, such as saving money for the future (Hershfield et al., 2011), hold in the context of commitment contracts and smartphone usage reduction.

However, if one were to look at the single effect sizes between the four means of the subsamples *treatment-high-FSC*, *treatment-low-FSC*, *control-high-FSC*, and *control-low-FSC*, one can observe interesting effect sizes, which indicate that potentially significant results could be found for larger sample sizes. Thus, this investigation may still serve as a starting point for future research to investigate such a relationship. With that lens, the aforementioned differing results from previous literature (suggesting positive relationships between the concept of future-self-continuity and other self-control issues) may also prove to be coinciding by future explorations of this relationship or similar relationships between commitment contracts and other psychological concepts.

Even though not significant, the results could be interpreted as pointing towards the direction suggested in hypothesis two, i.e. that commitment contracts may have a stronger impact on individuals with a low future-self-continuity. This potentially in the future to be found effect would appear to largely stem from individuals with a high future-self-continuity being better at reducing their smartphone usage and achieving their goals when exposed to simple goal setting, i.e. without the use of commitment contracts. Halpern et al. (2012) suggested it is still to be investigated which populations should be addressed with commitment contracts, and that "contracts must be designed so that they do not simply attract highly motivated people who might change their behaviour on their own anyway" (Halpern et al., 2012, p. e522). This study's initiative to investigate psychological concepts in general and the

concept of future-self-continuity in specific, may have provided a first indication for future research to find better answers to this question still to be explored.

5.3 Qualitative add-on: A quick look at good deeds performed

Lastly, even though not relevant for any existing theories or discussions in the research field of commitment contracts, this section shall give a quick overview of the good deeds that were performed in this study for not achieving daily goals. Generally, 66% of the participants exposed to commitment contracts reported good deeds for failing daily screen time goals. The remaining 13 participants reported “no deed”. Performed good deeds ranged from small, seemingly irrelevant actions, to larger and unexpected actions. Table 7 presents a few but not all examples.

Table 7 Examples of good deeds performed

“Day 1: no fail. Day 2: I made a \$5 donation to my favorite podcast creator. Day 3: no deed. Day 4: no deed. Day 5: successfully stayed under limit. Day 6: offered to explain the system for income tax to a colleague (recent immigrant and foreign national) and help file taxes for the year. Day 7: successfully stayed under limit. “
“I helped a colleague with a task, that was not my responsibility at all.”
“Bought tacos for a homeless person.”
“Gave money to a homeless person”
“I made my partner his favorite dessert.”
“I refilled a water bottle, I cooked food for friends, I invited friend for dinner, I offered to do the dishes.”
“1. Say compliments to friends and relatives (5 times) 2. Help with tasks (1 time) 3. Buy “Situation Stockholm”, a magazine for the benefit of the homeless (1 time)”
“I helped my friend with her cv and cover letter - I made my boyfriend breakfast - I did my boyfriend’s laundry - I recruited two more participants for this experiment - I gave my sister a few pieces of clothes that I know she likes and I do not really use anymore - I bought my mom a nice gift.”

The good deeds performed shall not be evaluated in connection to any quantitative inquiries of this experiment or commitment contract effectiveness in general. Rather, they shall serve as examples of how the design of soft commitment contracts can have positive impacts on people around us and society as a whole. Instead of inducing a pure cost to the individual engaging in a commitment contract – such as in the case of hard commitment contracts where one loses money – this approach established a psychological cost of having to do something, which in turn comes with positive effects to others and potentially to the person doing the good deed herself. For instance, one subject from the Prolific Academics sample (which was not used for the quantitative analyses), reported that she “went to the hairdresser, and donated my hair to an organisation that makes wigs for children with cancer.” She further states this was something she had been thinking about before, but put off doing it for months. Thus, one can assume that this subject also experienced some utility and happiness from performing a good deed. This can be seen as an example of how the used design of soft commitment contracts can have positive impacts for other people (a little girl with cancer getting a wig with real hair) and to the person performing the good deed herself.

6. Conclusions

In the following section, key findings of this thesis will be presented and put into perspective. The first part presents a summary of the conducted study and consequent conclusions that can be made. Thereafter, theoretical and practical contributions based on these findings will be discussed. Lastly, this thesis will come to an end via a discussion of important and relevant limitations of the study as well as suggestions for future research.

The intention behind this master thesis was to investigate the effectiveness of commitment contracts in the context of smartphone usage, and further to find out whether individuals' future-self-continuity has an impact on that effectiveness. The apparent problems stemming from increasing smartphone usage in our society have attracted attention from media, practitioners, and researchers. However, researchers have only looked at smartphone usage from medical and mental health perspectives. Potential ways to initiate behavioural change have not been studied in this context. Because the effectiveness of commitment contracts in addressing issues of self-control and changing behaviour has been proven in other contexts, this study set out to explore whether similar positive results could be observed when addressing the issue of smartphone usage. Furthermore, a research gap was made out, arguing that the current field of commitment contract research has not looked at the interplay between any psychological concept and commitment contract effectiveness. This thesis aimed to contribute to the field by investigating such a relationship with individuals' future-self-continuity – a concept that has helped explain many other behaviours where self-control problems can arise. Concluding, this thesis intended to investigate whether commitment contracts could help people change their phone usage behaviours and whether said effectiveness varies depending on people's connectedness to their future selves.

In that, theories from the two different research fields were combined to design and conduct an experiment in which participants set out to decrease the time they spend on their phones for one week. The results of this study partially support the hypothesis that commitment contracts can have positive effects on smartphone usage behaviours. While the results cannot confirm that commitment contracts significantly reduce relative smartphone usage, they do suggest that commitment contracts help people in achieving daily screen time goals. Whether inconsistency between those closely connected conceptualizations of outcome variables comes from stronger fall-back behaviours (i.e. excessively more usage once a limit is surpassed) by individuals who achieve their daily goals more often, cannot be answered by this thesis and may be subject to future research. The results of this study further provide no support for the hypothesized negative moderating effect of future-self-continuity on commitment contract effectiveness. However, due to limited statistical power in this analysis, it can be concluded that a need for further empirical investigations of this relationship persists.

In summary, this study provides first answers to its research question, suggesting that commitment contracts have partial positive effects on individuals' attempts of reducing their smartphone usage, and not suggesting individuals' future-self-continuity levels differentially effecting these outcomes.

6.1 Theoretical contributions

Based on the conclusions that can be drawn from this study as outlined above, this thesis contributes both to the research field of commitment contracts and to the research field of future-self-continuity.

One main contribution to the growing body of research about commitment contracts is the transfer of a previously established solution to issues of self-control into a new context. As such, this thesis serves as a first attempt to expand the use of commitment contracts from classical self-destructive behaviours to the modern-day issue of smartphone usage. The partial

support for the effectiveness of commitment contracts in the context of smartphone usage lays the groundwork for further empirical explorations in this new context and should motivate future research to adopt, apply, and research commitment contracts in other unexplored contexts. Furthermore, by using a soft commitment contract, this thesis contributes to the scarce empirical field of such soft commitments. On a purely hypothetical note, this thesis also suggests that the design of a soft commitment may allow for positive side-effects for individuals engaging in such contracts and other societal actors around them, as suggested to stem from good deeds in this thesis. However, future quantitative and qualitative research on this matter is needed to allow for better evaluation of mentioned potentially positive side-effects. Lastly, this thesis contributes to the research body of commitment contracts by adding a psychological layer to pursue the question of which individuals can be helped most by commitment contracts. In investigating the relationship between commitment contract effectiveness and individuals' future-self-continuity, the results do not allow to establish an explanatory relationship. However, this study serves as a first investigation of any relationship between individuals' personality characteristics and commitment contracts effectiveness. As such, it may inspire future investigations on this or any other relationship commitment contract effectiveness may have with psychological concepts.

Regarding the growing body of research on future-self-continuity, this thesis contributed to the theoretical field by investigation the concept in relation to a new self-control issue – smartphone usage. Even though no significant relationship could be established, this thesis applied the concept in a new context and thus expanded the knowledge about where future-self-continuity may or may not be of explanatory value for human behaviour.

6.2 Practical contributions

In addition to those theoretical contributions, this study also provides the basis for clear implications for various practitioners addressing people's issues of self-control. Specifically,

app developers, smartphone producers, or other actors who address the issue of smartphone usage reduction may learn a lot from this study's results. The finding that commitment contracts (partially) helped people change their phone usage behaviour may provide a promising basis for mechanisms and tools to be developed that can effectively shape human behaviour in a positive way. Furthermore, while the results of this study should not be blindly transferred into other contexts, one may argue that many practitioners, addressing people's self-control issues in any context, may find inspiration in this thesis' results. Soft commitment contracts showed to be promising in helping individuals change their smartphone usage behaviours and may be valuable to test or at least consider in future solutions intending to help individuals change other behaviours and achieve other goals as well. Additionally, the innovative design of a soft commitment contract in this study may serve as a basis for future designs by practitioners. Especially the idea to design soft commitment in a way that leads to potential positive externalities (i.e. increasing utility for the person herself and people around her) may be picked up by practitioners who prefer such an option over monetary punishments. Good deeds as a cost of soft commitment could be an interesting alternative to classical monetary costs of hard commitments.

6.3 Limitations

In discussing conclusions and contributions of this thesis, several limitations of the study cannot be disregarded and are outlined in this section.

6.3.1 General limitations stemming from the research design

Starting with general limitations of this study, the nature of a framed field experiment comes with some implications regarding internal and external validity. Compared to a lab experiment, this study did not allow for as much control and filtering out of external effects.

However, this was a calculated drawback of pursuing an experiment with field context in commodity and task, thus aiming to increase external validity.

Another limitation worth addressing is the length of the study. In conducting a one-week experiment, no longer-term effects in behavioural change can be observed. Given more time to conduct an experiment, a longer experimental period and follow-up data checks after the experimental period could have provided more details about long-term changes in smartphone usage and about potential reverting behaviours after the expiration of commitment contracts. Furthermore, the short time horizon provided a limited basis for data. Due to technical reasons in accessing data, baseline phone usage could only be computed based on the last seven days prior to start of the experiment. Since smartphone usage is volatile, longer time horizons for measuring baseline usage and for measuring treatment effects (i.e. experimental period) would have increased internal validity.

This study was further limited in the statistical power of performed analyses due to its sample size. With a finisher-sample size of 80 participants, one has 80% power to detect an effect size, according to Cohen's d , of 0.63, at a 5% confidence level in a two-sided t-test. With a larger sample, it could potentially have been possible to find significant effects for tests where the given results suggest no significance but where effect sizes are promising though smaller than what can be found to be significant given the sample size.

The sample population itself imposes further limitations to this study's external validity. Participation was based on individuals' voluntary decision to take part in the experiment. It is possible that this lead to a selection bias so that participants of the study differ from the individuals who declined to participate in the first place. Also, a large part of the sample consists of students, which is not representative of the population the sample aims to represent. While these points imply limitations regarding the study's external validity, internal validity

should not have been affected by any of these issues due to randomized allocation of participants into treatment and control group.

6.3.2 Limitations stemming from the investigated behaviour and how it is measured

The choice for this study's investigated behaviour – smartphone usage – may itself cause some limitations in trying to find ways to help people overcome behavioural problems and dependencies on technological devices. While data of smartphone usage is measured in an objective way, there is a potential for subjects to use other devices to access time wasting apps or websites via tablets, computers, or TVs. Some individuals may even have multiple phones. This could distort the results and over-exaggerate the effectiveness of commitment contracts as participants in the treatment group may be incentivised to use other devices as a substitute when soft penalties for the use of their smartphones are possible. That is especially true at the margin where penalties are imposed. This limitation is very difficult to get around due to technological reasons. For this study, the inconvenience of not being able to observe substitutional behaviours had to be taken as a given.

Also, self-reporting of screen times could be a source of dishonesty in participants' recorded screen time data. This could have had an impact on the results if individuals in the treatment group were somehow affected by the commitment contract to be more prone to report untruthful data. In this sense, this thesis was relying on the good faith of the volunteers.

Continuing with limitations due to the choice of investigating smartphone usage, the variable as such showed to have a lot of variance. In that, daily smartphone usage times are a difficult measure to be used. It can be affected much more by external forces (e.g. weather, time of the year/month/week/, current workload) than for instance weight loss. This has impacts on relative screen time reduction as a variable. Large variations in this variable could be

observed. Weight loss or binary variables are a lot more stable and centered, which makes it easier to find significant effects for small sample sizes. This could also be a reason for why findings for hypothesis 1b) are significant and findings for hypothesis 1a) are not. The standard deviations, relative to the observed means, are higher for relative smartphone usage reduction than for daily goals achieved. Furthermore, smartphone usage's volatility also impacts the outcome variable daily goals achieved. Due to its binary nature, this variable misses out on detailed behaviour. For instance, it does not consider how much an individual surpasses a goal.

6.3.3 Limitations stemming from choice and design of the commitment contract

This leads us to limitations due to the choice and design of the soft commitment contract used in this study. Ultimately, an either-or penalty was used. Once a limit was surpassed, participants had no incentive to not use their phone excessively much. It could be that this had negative impacts on the relative reduction variable. Stepwise costs could address this issue in the future. Furthermore, the application of good deeds to be performed as a penalty is more difficult to enforce than for instance a deposit contract, where individuals give away money beforehand and only receive it back if a certain goal is met. For this study, enforcement of the penalty hinged upon the honesty and integrity of participants.

6.3.4 Other limitations

Furthermore, the choice of how to utilize the concept of future-self-continuity as an explanatory variable for commitment contract effectiveness comes with limitations. Especially the conducted median split, categorizing participants as High or Low FSC individuals comes with certain drawbacks. First, such a split implies the assumption of a linear relationship between this variable and the effects it tries to explain. This limits the analysis and does not

allow for investigation of any relationship other than a linear one. Second, transforming a continuous variable into a categorical variable reduces the observed sizes of the effects. (Allen, 2017)

In trying to establish control variables relevant for the studied behaviour of smartphone usage, no variable was included that controlled for participants' willingness or desire to decrease their smartphone usage. The non-inclusion of such a variable is simply rooted in that the researcher thought about such a variable too late when a large part of participants had already finished the experiment. This would have constituted an interesting control variable and should be included in and addressed by future research. Introducing this variable may also be relevant to put contributions for practitioners in better perspective, providing insights regarding who wants to decrease his/her screen time and how commitment contract effectiveness can be evaluated in relationship to those individuals.

As a last limitation, the problem of missing outcome data must be addressed. Participants who are lost to follow up, are a considerable issue for all intention-to-treat analyses. One cannot know whether dropouts in this experiment did not care about the study or decreasing their phone usage in general, whether they forgot to report their data, got frustrated during the process, or were hindered in any other way. Some participants might have encountered technical issues or were subject to external forces hindering them from finishing the experiment. The sensitivity analysis treating dropouts as having reverted to baseline is one approach of trying to deal with this problem, but comes with drawbacks itself, because any assumptions for outcome variables are strong and artificial. Missing outcome data is simply a difficult problem to deal with in applying intention-to-treat principles. (Alshurafa et al., 2012)

All analyses and results should be interpreted with respect to the various limitations outlined above. Limitations referring to the study design as well as variable conceptualizations may have partially made it difficult to find significant results for the proposed hypotheses.

Furthermore, the limitations have an impact on this study's reliability and internal or external validity, which should not be disregarded in evaluating results or generalizing findings.

6.4 Future research

One issue to be addressed by future research is the limited length of this study. Longer-lasting experiments are needed to establish insights regarding long-term effects in behavioural change. Such research could also observe potential reverting smartphone usage behaviours after the expiration of commitment contracts. It would be interesting to look at such residual effects. Previous research on other self-control issues suggests that subjects would revert to normal behaviour. While one could assume to observe the same for phone usage, no positive statements can be made without future research about such effects.

Furthermore, the field of commitment contracts is still in need of experiments with larger sample sizes that support or expand this study's findings regarding the effectiveness of commitment contracts in helping individuals to decrease the time they spend on their phones. Such studies will increase statistical power of analyses and allow to make stronger suggestions about hypothesized effects. Additionally, conducting experiments for which the sample may be more representative of the population it aims to represent are in need to generalize findings for the population.

In addressing the general societal issue of increasing screen times and dependencies on technological devices such as smartphones but also TVs, computers, or tablets, future research may try to absorb and observe substitutional behaviours.

To establish commonly accepted measures for future research, the outcome variables and measures created and used in this thesis, capturing commitment contract effectiveness in the context of smartphone usage, need further validation. Testing different alterations may allow to find the most suitable measures for future experiments. The same applies for the design of

the soft commitment contract used in this study. Furthermore, effectiveness investigations of different types of commitment contracts in the context of smartphone usage are needed to compare which types of commitments may serve best in altering smartphone usage behaviours. As such, future research could consider the use of stepwise contracts, where there is not just one penalty for surpassing a daily limit but continuous penalties, so that for instance every other fifteen minutes of surpassing a limit, the cost increases. Also, the use of deposit contracts or other hard commitments could be investigated in the future. Hybrid forms of commitment contracts combining hard and soft commitments may also be subject to such investigations.

With the goal to generate insights regarding which individuals may benefit the most from the use of commitment contracts, future research should re-investigate the proposed moderating effect of future-self-continuity in high power studies. Such efforts may further be complemented by exploring the relationship between commitment contract effectiveness and other psychological constructs that appear to be promising explanatory variables in theory.

Given the general findings and limitations of this thesis as well as the suggestions for future research, one can see the great potential for future research to explore more aspects regarding commitment contracts in the context of smartphone usage. Furthermore, there is a multitude of opportunities to study commitment contracts in other contexts, with different commitment contract designs, in combination with alternative psychological concepts, in different settings, or with different samples. Such efforts will enrich the field of commitment contracts and increase our understanding of their effectiveness in the context of smartphone usage and other contexts. Ultimately, all related future research efforts will contribute to the tackling of self-control problems and help people to live up to their own goals.

7. References

Allen, M. (2017). *The sage encyclopedia of communication research methods* (Vols. 1-4).

Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781483381411.

Retrieved from: <https://methods.sagepub.com/Reference/the-sage-encyclopedia-of-communication-research-methods/i8685.xml>

Alshurafa, M., Briel, M., Akl, E., Haines, T., Moayyedi, P., Gentles, S., Rios, L., Tran, C.,

Bhatnagar, N., Lamontagne, F., Walter, S., & Guyatt, G. (2012). Inconsistent Definitions for Intention-To-Treat in Relation to Missing Outcome Data: Systematic Review of the Methods Literature. *PLoS ONE*, 7(11), e49163. doi: 10.1371/journal.pone.0049163.

Retrieved from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3499557/>

Anderson, M. & Silver, L. (2019). *7 key findings about mobile phone and social media use in emerging economies*. [online] Pew Research Center. Retrieved from:

<https://www.pewresearch.org/fact-tank/2019/03/07/7-key-findings-about-mobile-phone-and-social-media-use-in-emerging-economies/>

Andreassen, C., Pallesen, S., & Griffiths, M. (2017). The relationship between addictive use of social media, narcissism, and self-esteem: Findings from a large national survey. *Addictive Behaviors*, 64, pp.287-293.

Angrist J., Lang D., & Oreopoulos P. (2009). Incentives and Services for College

Achievement: Evidence from a Randomized Trial. *American Economic Journal: Applied Economics*, 1(1), pp.136-163.

- Angrist J., & Lavy V. (2009). The Effects of High Stakes High School Achievement Awards: Evidence from a Randomized Trial. *American Economic Review*, 99(4) pp.1384–1414.
- Ashraf N., Karlan D., & Yin W. (2006). Tying Odysseus to the Mast: Evidence From a Commitment Savings Product in the Philippines. *The Quarterly Journal of Economics*, 121 (2), pp.635–672.
- Barrick, M. & Mount, M. (1991). The Big Five Personality Dimensions and Job Performance: A Meta-Analysis. *Personnel Psychology*, 44(1), pp.1-26.
- Bartels, D. & Rips, L. (2010). Psychological connectedness and intertemporal choice. *Journal of Experimental Psychology: General*, 139(1), pp.49-69.
- Blouin-Hudon, E. & Pychyl, T. (2015). Experiencing the temporally extended self: Initial support for the role of affective states, vivid mental imagery, and future self-continuity in the prediction of academic procrastination. *Personality and Individual Differences*, 86, pp.50-56.
- Bryan, G., Karlan, D., & Nelson, S. (2010). Commitment Devices. *Annual Review of Economics*, 2(1), pp.671-698.
- Bryman, A. & Bell, E. (2015). *Business research methods*. (4.ed.). Oxford: Oxford University Press.

Charness G., & Gneezy U. (2009). Incentives to Exercise. *Econometrica*, 77(3), 909-931.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. (4.ed.). Thousand Oaks, CA: SAGE Publications, Inc.

Cui, X., Shi, Y., & Xu, L. (2017). Alleviating Time Inconsistent Behaviors via a Competition Scheme. *Naval Research Logistics (NRL)*, 64(5), pp.357-372.

Della Vigna S., & Malmendier U. (2006). Paying Not to Go to the Gym. *American Economic Review*, 96(3), pp.694-719.

Deloitte (2017). *State of the smart – Global Mobile Consumer Survey 2017: UK Cut*.

[online] London: Deloitte. Retrieved from:

https://www.deloitte.co.uk/mobileuk2017/assets/img/download/global-mobile-consumer-survey-2017_uk-cut.pdf

Deloitte (2018). *Deloitte survey: Smartphones continue to reign supreme as consumers'*

preferred device. [online] New York: Deloitte. Retrieved from:

<https://www2.deloitte.com/us/en/pages/about-deloitte/articles/press-releases/deloitte-launches-2018-global-mobile-consumer-survey.html>

Dudovskiy, J. (2016). *The Ultimate Guide to Writing a Dissertation in Business Studies: A Step-by-Step Assistance*. (1.ed). [e-book]. Research-methodology.net

- Falk A., and Kosfeld M. (2006). The Hidden Costs of Control. *American Economic Review*, 96(5), pp.1611-1630.
- Fitzsimons, G., & Finkel, E. (2010) Interpersonal Influences on Self-Regulation, *Current Directions in Psychological Science*, 19(2), pp.101-105.
- Giné X., Goldberg J., & Yang D. (2010). Identification Strategy: A Field Experiment on Dynamic Incentives in Rural Credit Markets. *SSRN Electronic Journal*.
doi: 10.2139/ssrn.1633517.
- Gino, F., Schweitzer, M., Mead, N. & Ariely, D. (2011). Unable to resist temptation: How self-control depletion promotes unethical behavior. *Organizational Behavior and Human Decision Processes*, 115(2), pp.191-203.
- Halpern S., Asch D., & Volpp K. (2012). Commitment contracts as a way to health. *BMJ*, 344, pp.e522-e522. doi: 10.1136/bmj.e522. Retrieved from:
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4707874/>
- Halpern, S., French, B., Small, D., Saulsgiver, K., Harhay, M., Audrain-McGovern, J., Loewenstein, G., Brennan, T., Asch, D. & Volpp, K. (2015). Randomized Trial of Four Financial-Incentive Programs for Smoking Cessation. *New England Journal of Medicine*, 372(22), pp.2108-2117.
- Hardin, J. (1998). The Principle of Intention to Treat Analysis in Clinical Studies: Its Use and Controversies. *Journal of Pharmacy Practice*, 11(4), pp.231-238.

- Harrison, G. & List, J. (2004). Field Experiments. *Journal of Economic Literature*, 42(4), pp.1009-1055.
- Heritier S., Gebski, V., & Keech, A. (2003). Inclusion of patients in clinical trial analysis: the intention-to-treat-principle. *MJA*, 179, pp.438-440.
- Hershfield H., Garton M., Ballard K., Samanez-Larkin G., & Knutson B. (2009). Don't stop thinking about tomorrow: Individual differences in future self-continuity account for saving. *Judgment and Decision Making*, 4(4), pp.280–286.
- Hershfield, H. (2011). Future self-continuity: how conceptions of the future self transform intertemporal choice. *Annals of the New York Academy of Sciences*, 1235(1), pp.30-43.
- John L., Loewenstein G., Troxel A., Norton L., Fassbender J., & Volpp K. (2011). Financial Incentives for Extended Weight Loss: A Randomized, Controlled Trial. *Journal of General Internal Medicine*, 26(6), pp.621-626
- Kaur S., Kremer M., & Mullainathan S. (2015). Self-Control at Work. *Journal of Political Economy*, 123(6), pp.1227-1277.
- Kemp, S. (2018). *Digital in 2018* [online] We Are Social. Retrieved from: <https://wearesocial.com/blog/2018/01/global-digital-report-2018>

Kremer M., Miguel E., & Thornton R. (2009). Incentives to Learn. *Review of Economics and Statistics*, 91, pp.437- 456.

Laibson D. (1997). Golden Eggs and Hyperbolic Discounting. *The Quarterly Journal of Economics*, 112(2), pp. 443-477.

Levitt S., List J., Neckermann S., & Sadoff S. (2016). The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance. *American Economic Journal: Economic Policy*, 8(4), pp.183-219.

Mertler, C. (2016). *Introduction to Educational Research*. Thousand Oaks, CA: SAGE Publications, Inc.

O'Donoghue T., & Rabin M. (1999). Doing It Now or Later, *American Economic Review*, 89(1), pp.103-124.

O'Donoghue T., & Rabin M. (2002). *Addiction and Present-Biased Preferences*. Berkely, CA: Institute of Business and Economic Research.

Patterson R. (2015). Can Behavioral Tools Improve Online Student Outcomes? Experimental Evidence from a Massive Open Online Course. *Journal of Economic Behavior & Organization*, 153, pp.293-321.

Perdue, B., & Summers, J. (1986). Checking the success of manipulations in marketing experiments. *Journal of Marketing Research*, 23 (4), pp.317-326.

- Rogers T., Milkman K., & Volpp K. (2014). Commitment Devices: Using Initiatives to Change Behavior. *JAMA*, 311(20), pp.2065-2066.
- Rothmann, S. & Coetzer, E. (2003). The big five personality dimensions and job performance. *SA Journal of Industrial Psychology*, 29(1), pp.68-74.
- Saunders, M., Lewis, P., & Thornhill, A. (2012). *Research methods for business students*. (6.ed.) Harlow: Pears.
- Schwarzer, R. & Fuchs, R. (1996). Self-efficacy and health behaviors. In Conner M. & Norman P. (1995). *Predicting health behavior: Research and practice with social cognition models*. pp. 163-196. Buckingham, UK: Open University Press.
- Stedman-Falls, L., Dallery, J., & Salloum, R. (2018). Discrete Choice Experiments on The Acceptability of Monetary-Based Health Treatments: A Replication and Extension to Deposit Contracts. *The Psychological Record*, 68(4), pp.501-512.
- Strotz, R. (1955). Myopia and Inconsistency in Dynamic Utility Maximization. *The Review of Economic Studies*, 23(3), pp. 165-180.
- TEDx Talks. (2011). *Self control: Dan Ariely at TEDxDuke* [Video file]. Retrieved from: <https://www.youtube.com/watch?v=PPQhj6ktYSo>

- Thaler, R. & Benartzi, S. (2004). Save More Tomorrow™: Using Behavioral Economics to Increase Employee Saving. *Journal of Political Economy*, 112(1), pp.164-187.
- Thorgeirsson T., & Kawachi I. (2013). Behavioral Economics: Merging Psychology and Economics for Lifestyle Interventions. *American Journal of Preventive Medicine*. 44(2), pp.185-189.
- Twenge, J., Joiner, T., Rogers, M., & Martin, G. (2017). Increases in Depressive Symptoms, Suicide-Related Outcomes, and Suicide Rates Among U.S. Adolescents After 2010 and Links to Increased New Media Screen Time. *Clinical Psychological Science*, 6(1), pp.3-17.
- van Gelder, J., Hershfield, H. & Nordgren, L. (2013). Vividness of the Future Self Predicts Delinquency. *Psychological Science*, 24(6), pp.974-980.
- van Gelder, J., Luciano, E., Weulen Kranenbarg, M., & Hershfield, H. (2015). Friends with My Future Self: Longitudinal Vividness Intervention Reduces Delinquency. *Criminology*, 53(2), pp.158-179.
- Volpp, K., John, L., Troxel, A., Norton, L., Fassbender, J., & Loewenstein, G. (2008). Financial Incentive–Based Approaches for Weight Loss. *JAMA*, 300(22), p.2631-2637.

Volpp K., Troxel A., Pauly M., Glick H., Puig A., Asch D., Galvin R., Zhu J., Wan F., DeGuzman J., Corbett E., and Weiner J. (2009). A Randomized, Controlled Trial of Financial Incentives for Smoking Cessation. *New England Journal of Medicine*, 360(7), pp.699-709.

White, I., Horton, N., Carpenter, J., & Pocock, S. (2011). Strategy for intention to treat analysis in randomised trials with missing outcome data. *BMJ*, 342, pp.d40-d40.
doi: <https://doi.org/10.1136/bmj.d40>. Retrieved from:
<https://www.bmj.com/content/342/bmj.d40>

8. Appendix

Appendix 1 – Results from pre-study 1

Table 8. Baseline characteristics; iPhone vs Android users

	Phone Type	
	iPhone n=32	Android n=28
Gender		
Male	19 (59.4)	19 (67.9)
Female	13 (40.6)	9 (32.1)
Other	0	0
Age (sd)	25.22 (3.86)	28.36 (11.16)
Marital status		
Single	25 (78.1)	22 (78.6)
Married	3 (9.4)	3 (10.7)
Separated	0	0
Divorced/widowed	1 (3.1)	1 (3.6)
Don't want to disclose	3 (9.4)	2 (7.1)
Employment status		
Student	22 (68.8)	16 (57.1)
Employed	9 (28.1)	7 (25)
Self-employed	1 (3.1)	2 (7.1)
Unemployed	0	0
Other	0	0
Working hours (sd)	34.2 (15.7)	31.9 (15.4)
FSC (sd)	5.05 (0.78)	4.88 (0.92)

Note: numbers in brackets indicate percentages within the given group unless standard deviation is indicated by (sd)

Appendix 2 – PDF version of Survey 1 as used for pre-study 2 and the main study

Block 1: Introduction

Welcome to the Phone Usage Experiment!

You will participate in a master thesis experiment at Stockholm School of Economics.

For the next week, you should try to decrease your daily smartphone usage.

The study consists of 3 parts and will take 10-15 minutes in total.

Part 1: Survey 1 (this survey) – takes 3-5 minutes

Part 2: Experimental week – decrease your daily smartphone usage for 7 days – noting down usage times takes around 1 minute/day

Part 3: Survey 2 – takes 2-4 minutes

Your identity and behavior in this study will be anonymous to other participants. Please do not discuss this experiment with others during or after your participation.

Reminder: You can only participate in this experiment if you use an iPhone.

Online Survey Consent Form

Purpose of research:

- 1) Examining ways to decrease people's phone usage.
- 2) Studying people's phone usage connected to personal characteristics such as their future-self-continuity.

Time required:

The experiment will go on for 1 week. You are asked to answer one survey today, and another survey after the experimental period. Participation will take 10-15 minutes in total.

Risks: There are no known risks associated with this research study. Your participation in this study will remain anonymous to other participants.

Benefits: 1 out of every 15 participants who take part throughout the whole experiment will be randomly selected to receive 200 SEK / 20€.

Participation and withdrawal: Your participation in this study is completely voluntary and you can withdraw at any time.

Researcher contact: If you have questions about this research, please contact Max Valentin Bartels from the Stockholm School of Economics at 41193@student.hhs.se

By clicking the next-button you state that you have read and understood this consent form and agree to participate in this research study.

Block 2: Future self-continuity

Future self-continuity

One aspect of this study investigates the relationship between people's connectedness to their future self and their phone usage behavior.

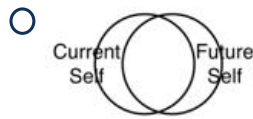
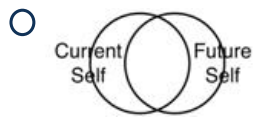
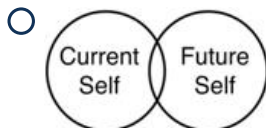
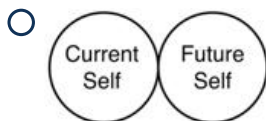
The following questions measure your "future self-continuity".

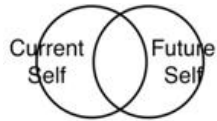
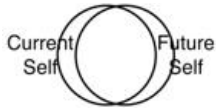
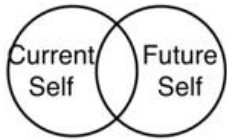
Please try to imagine yourself 10 years from now and answer based on your instinct.

Current self = you right now

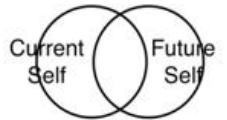
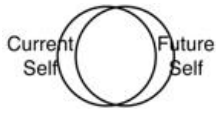
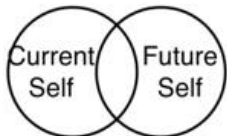
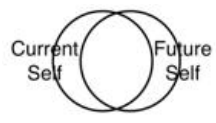
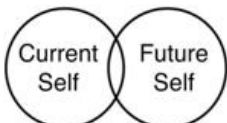
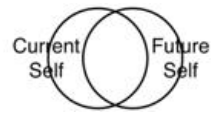
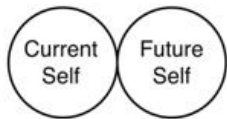
Future self = you in the future (ten years from now)

Select the circle pair that best describes **how similar** you feel to your future self ten years from now





Select the circle pair that best describes **how connected** you feel to your future self ten years from now



How much do you care about your future self ten years from now?

Don't care at all

A moderate amount

Care a great deal

- 3
- 4
- 5

Gender

- Female
- Male
- Other
- Don't want to disclose

Age

Marital status

- Single
- In a relationship
- Married
- Separated
- Divorced / Widowed
- Don't want to disclose

E-Mail address

(to send a link for Survey 2. You won't receive any spam.)

Block 4: Phone Usage

What is the average daily time you spent on your smartphone in the last week?
(Please indicate in minutes rounded up)

Example: 2 h 17 min = 137 min

(in your iPhone, go to "Settings" - "Screen Time" - "my iPhone" - "Last 7 days")

Throughout the next week, you should set yourself the goal of decreasing your daily smartphone time by 30%.

Please indicate your new daily screen time goal below (in minutes)

Calculate: previously reported average screen time x 0.7

Example: 137 min x 0.7 = 96 min

Important: Please make sure you remember your new daily screen time goal! You should have it in mind for the next 7 days.

Please set your daily screen time goal in your iPhone.

In your iPhone, go to "Settings" - "Screen Time" - "App Limits" - "Add Limit" - "All Apps & Categories" - set your agreed upon daily screen time goal.

Your phone will notify you 5 minutes before you reach your daily limit. This will not shut down your apps once you surpass that limit. You will still be able to use your phone normally.

Starting tomorrow morning, please write down your daily screen times on a piece of paper.

Write your daily goal on top of the paper. Then add your actual screen time of the previous day every morning. Set a daily alarm for the morning so that you do not forget to write down your actual screen times.

Block 5: Treatment Condition

Commitment contract

One part of this study is to investigate the helpfulness of commitment contracts on smartphone usage.

A commitment contract is a contract you make with yourself to achieve your future goals.

- 1) You commit to a certain behavior in the future
- 2) You commit to a cost in case you do not follow up on your planned behavior.

This mechanism increases the motivation to act upon your goals.

To help you achieve your daily goals in this experiment, please sign this "commitment contract":

**"I intend to stay below my daily screen time goal for 7 days, starting today.
For every day that I fail to do so, I will do something good for someone else on the following day.**

I will note down my good deeds to report them in Survey 2."

Please sign below.



A rectangular box containing the text "SIGN HERE" in large, bold, grey capital letters. Below the text is a horizontal line. To the left of the line is a small "x" icon, and to the right is a red "clear" button.

Next steps:

- Starting today, try to stay below your maximum daily screen time for 7 days.
- Starting tomorrow, check your actual screen time of the previous day in your phone settings.

Write it down on a piece of paper. You should do this every morning.

Example: Tomorrow morning you write down today's actual screen time.

(Set a daily alarm, so you don't forget to do this every day).

- Remember your commitment contract!

If you fail to stay below a daily goal, do something good for someone else and note down the good deed on the same piece of paper.

- In one week, you will be sent a link for Survey 2.

Fill out Survey 2 to report your data and answer a few follow-up questions.

If you have any questions regarding the experiment, please reach out to 41193@student.hhs.se

Please click the next button to submit Survey 1!

Block 5: Control Condition

Goal Setting

Please set yourself this goal:

"I will try to stay below my new daily screen time goal for 7 days, starting today."

There are no formal penalties or rewards for achieving your new daily screen time goals.

Next steps:

- Starting today, try to stay below your maximum daily screen time for 7 days.

- Starting tomorrow, check your actual screen time of the previous day in your phone settings.
Write it down on a piece of paper. You should do this every morning.
Example: Tomorrow morning you write down today's actual screen time.
(Set a daily alarm, so you don't forget to do this every day).

- In one week, you will be sent a link for Survey 2.
Fill out Survey 2 to report your data and answer a few follow-up questions.

If you have any questions regarding the experiment, please reach out to 41193@student.hhs.se

Please click the next button to submit Survey 1!

Appendix 3 – PDF version of Survey 2 as used for pre-study 2 and the main study (treatment version)

Block 1: Introduction

Thank you for coming back!

This survey is the last part of this phone usage experiment. Answering will take approximately 2-4 minutes.

As before, your identity and your behavior in this study will be anonymous to other participants.

Please do not discuss this experiment with others after your participation.

Completion of this survey will automatically enter you into a lottery.

One out of 15 participants will be randomly selected to receive 20€ / 200 SEK.

If you have questions about this research, please contact Max Valentin Bartels from the Stockholm School of Economics at 41193@student.hhs.se

Block 2: Phone usage during experimental week

What was your agreed upon daily screen time goal based on Survey 1 ? (in minutes)

Please click on 2 for us to make sure that you are actually reading this

What was your actual screen time for every single day of the experiment? (in minutes)

(in your iPhone, go to "Settings" - "Screen Time" - "my iPhone" - "Last 7 days")
hold your finger on the bars to see exact daily screen time

Day 1	<input type="text"/>
Day 2	<input type="text"/>
Day 3	<input type="text"/>
Day 4	<input type="text"/>
Day 5	<input type="text"/>
Day 6	<input type="text"/>
Day 7	<input type="text"/>

Briefly name all the good deeds you did for every day that you failed to achieve your goal.

If you never failed, please write "no fail".

If you did not perform good deeds even though you failed your daily goal at least once, write "no deed".

Please enter your e-mail address
(to match your data from Survey 1 & 2)

Powered by Qualtrics