UNIVERSITY OF AGDER

Can nudging reduce default rates?

A field experiment testing the effect of nudging on young adults' propensity to default on cell phone bills.

Jorunn Lie Bjelland
and
Erlend Stene

Supervisor
Ellen Katrine Nyhus

This Master's Thesis is carried out as a part of the education at the University of Agder and is therefore approved as a part of this education. However, this does not imply that the University answers for the methods that are used or the conclusions that are drawn.

Abstract

Can altering the context in which people make their decisions be an effective method for reducing default rates? This thesis reports the results of an experiment that tested the effect of so-called «nudging» on young adults' decision-making process concerning bill payments. Based on theories of judgment and decision-making, text messages were sent out to 2500 randomized treatment and control subjects aged 18-30, pulled from One Call's (Network Norway AS) customer base. Incoming payments were then recorded to determine whether the messages had influenced customers' decision to pay their bill within due. The results provide evidence that nudging may positively influence people's decisions, and subsequently reduce default rates, if the nudge is executed correctly.

The aim of this study has been to test existing theories and models in this field of research in a new context, thereby contributing to further development and hopefully active use of nudging as a positive alternative to regulations. This could benefit each individual being nudged, and ultimately bring out positive larger scale societal changes.

Our hope is that companies that deal with payment collection will benefit from this research. Although former studies have shown the effect of nudging in different contexts, the results might not be directly transferrable to payment rates. We have therefore opted to empirically test the effect of nudging in a new context; on customers' payment behavior.

Key words: Behavioral economics, nudging, personal finances, young adults, field experiment, default rate, bill payments, cell phone industry.

Preface

This Master's Thesis is written as a final part of the Master`s Programme in Business Administration at University of Agder, Kristiansand School of Management. The thesis addresses the concept of nudging, a term belonging to the field of behavioral economics, which lies in the gap between economics and psychology. Behavioral economics has been a central part of several subjects during our Master's degree, particularly in the course ORG-419: Judgment and Decision-making, which we both attended last fall. It is especially this course that sparked our interest for the field and a desire to further explore the concept of nudging.

Writing a thesis is a demanding process. Through the process we have been challenged both academically, personally and in terms of team work. Overcoming these challenges has made us better equipped to face the workplace we are at the verge of entering. In spite of the tight schedule, we recognized the need to take time to wonder and let the information sink in, to fully capture the essence of our project and how to express ourselves accurately in this thesis. We hope that this is reflected in the final thesis.

Conducting a field experiment in such a limited time period carried many risks. Had we fully grasped the extent of these risks and the challenges ahead before we started this project, it might have never happened. Luckily we were blissfully unaware of the frustration to come, and we are very pleased with the overall process.

We would like to use this opportunity to express our gratitude and appreciation to several people for their invaluable help and support. First and foremost; we thank our supervisor Professor Ellen Katrine Nyhus for her enthusiasm, constructive feedback and support throughout the process. Her knowledge of and dedication to this field of research has been a strong asset in the process. Further, we thank Frode Elverum, Per Ola Stålberg and Bente Gjertsen at One Call (Network Norway AS) for believing on our project and making the necessary resources available for the implementation. We also thank Associate Professor Rotem Shneor and PhD Research Fellow Andrew Muteti Musau for valuable input and discussions, and Research Librarian Henry Langseth for technical support.

Lastly we would like to thank our professors and fellow students for a great time here at the University of Agder, and each other for a good collaboration.

Kristiansand, June 3rd 2013

Jorunn Lie Bjelland and Erlend Stene

Contents

List of figures

List of tables

1. Introduction

The purpose of this study is to investigate whether changing the context in which people make their decisions may have an effect on, and subsequently reduce, people's propensity to default payments. The study is motivated by an increasing concern for the rising debt levels in Norway, especially for young adults.

Current Western societies are often labeled as being excessively regulated. Some say that the law already interferes with individuals' freedom of choice, micromanaging the private sphere. As formal laws have limited bounds concerning the regulation of private life, researchers recognize a need for alternative ways to encourage people to make better choices for themselves, ultimately benefitting society as a whole. We see the potential in nudging as a positive alternative to more intervening forms of regulations, so long that it is used exclusively for positive purposes.

According to classical economic theory, all participants in the marketplace are assumed to be rational actors. By studying the effects of nudging, we investigate the implications of modifying this standard economic assumption, examining the possibility that actors in the economy deviate from the rational standard. «Nudging» is a relatively new term, first coined by Thaler and Sunstein in 2008. But the interest for research on this phenomenon has rapidly gained momentum; David Cameron and Barack Obama are big advocates of nudge theory, and it is catching on in companies all over the world.

The media fronts a growing concern for the state of young adults in Norway regarding their personal finances. Norwegian collection agencies and other institutions express a well-grounded concern for the increasing number of young adults that find themselves in financial difficulties due to defaulted payments and debt problems. We wanted to explore whether nudging might help change this negative development in personal finances for young adults. The primary reason for examining the effects of nudging on young adults' propensity to default on bill payments is that we have not been able to uncover any previous studies performed concerning this particular issue. As cell phone bills are highlighted in the media as bills that are typically defaulted on, we have focused our thesis on these particular payments.

With our thesis we aim to contribute to the growing field of research on how people's choices can be influenced in beneficial ways. We want to test the psychological tool of nudging in a context where it has not to our knowledge been tested before; on actual payment rates. Our goal is to lower the number of young adults that default their bills, serving a sociopolitical purpose as well as the purpose of bettering personal finances for individuals. The study was designed to answer the following research question;

Can nudging young adults reduce their propensity to default on cell phone bills?

We start this thesis with an overview of the current state of personal finances of young adults in Norway, and the theoretical foundation that sparked our interest to further investigate the effects of nudging. We then move on to a review of relevant experiments, examining what studies in the field already have discovered, and possible weaknesses in the performed research. In the next chapter we address the specific type of methods used to perform this study, introduce the experiment and the procedure for its implementation. We then present the results of our study. Based on these findings, we argue that nudging may significantly lower default rates. The thesis closes with a discussion of our findings, the limitations and implications of the study, and conclusions that can be drawn on the basis of our experiment.

2. Background and theoretical foundation

2.1 Background

Recent studies and statistics have shown a significant increase in the amount of outstanding credit debt in Norway, particularly owed by young adults, raising valid cause for concern. Outstanding credit debt includes money owed to credit card companies, sums owed or defaulted on uncollateralized loans, and money owed to collection agencies. By the end of 2011, young adults in Norway between the age of 19 and 26 owed NOK 1,1 billion to the aforementioned institutions (Stenseng, 2012). This is an increase of NOK 200 million from the year before (Stenseng, 2012).

Brusdal and Berg (2011) recently published a report on young adults in Norway and their credit-financed consumption. The report was developed based on statistics from collecting- and credit agencies, and caused governmental concern and initiative to investigate these issues further. The report documented a sharp increase in the number of young adults who default on credit loans and bill payments. In 2011, one in four who defaulted payments to the extent of forced debt collection, was between 18-25 years old (Brusdal & Berg, 2011). Further, the number of young adults that were registered with reduced credit scores had increased by 50% since 2005 (Brusdal & Berg, 2011).

The report states that young adults is not the age group with the largest default problems. Adults ranging from 40 to 44 years of age is actually the age group with the most severe financial problems (Brusdal & Berg, 2011). Young adults is, however, the age group that accounts for the sharpest increase in default rates and more frequently end up having debt problems. We have chosen to focus on young adults due to the severity of this development, and also because of the detrimental effects that having poor credit rates may have later in life. Young adults with poor credit ratings will have a difficult time getting approved for mortgages, and the credit debt they build up in the initial phases of adult life may have crippling effects on their personal finances later on.

Brusdal and Berg's report discusses a variety of partially interconnected reasons for why so many young adults accumulate staggering amounts of credit debt and default on their bill payments. There are two factors that seem to be more prevalent than others: a lack of knowledge concerning money and contract management, and a level of consumption that is not sustainable with the

income and funds available to them (Brusdal & Berg, 2011). Young adults are entering a phase in life where they encounter new and unfamiliar situations. At this stage of the life-cycle, many are becoming financially independent for the first time, have just moved away from home, and many are students whose sole source of income are part-time jobs, student loans and scholarships. They sign contracts for housing, cell phone plans, internet and electricity, that all have to be managed and paid for. Most young adults have very limited knowledge and experience in this field, making them vulnerable for failure and exploitation if they are not fully grasping what they are getting into (Brusdal & Berg, 2011). At the same time, young adults have a relatively high consumption compared to their income, and when compared to previous generations and other stages of the life cycle, their social arena is highly commercialized. This can easily lead to unwise spending and consumption, and subsequently credit financed consumption due to a lack of sufficient funds (Brusdal & Berg, 2011).

Brusdal and Berg's report (2011) divides consumption into three categories; food and shelter as the most basic, consumption of social goods comes second, and lastly experience and pleasure consumption. Naturally, most of the 1000 respondents who were interviewed for this study listed basic goods as their first priority. Further, the authors uncovered that the majority of respondents reported that the consumption of social and pleasure goods was so important that they would feel poor if they were not able to extensively participate in the social arena (Brusdal & Berg, 2011). These goods and experiences were perceived to be so important that these young adults would to a large extent rather finance them with credit, than saving up the money beforehand and enjoy such goods and experiences less often. This attitude towards consumption may help explain why many young adults are defaulting on their bills and have substantial, and growing, credit debt (Brusdal & Berg, 2011).

Another contributing reason for why young adults are prone to wind up in financial difficulties is explained by Gulbrandsen (1999). In a report published by NOVA, Gulbrandsen proposes that defaulting on bill payments depends on the individual's willingness to pay, as he labels as the individual's «morality», as opposed to ability to pay. In his study, he found that timely bill payments are correlated with a person's morality, and that morality is highly related to age. Gulbrandsen suggests that morality is developed through the stages of the life cycle, and that people are therefore more likely to keep up with bill payments as they get older.

Further, Brusdal and Berg divide young adults into four categories based on the behavioral characteristics that got them into financial difficulties. These categories are set up along two dimensions; whether the problem is perceived to be caused by themselves or by external forces, and whether the individual made active or passive choices in the process.

Table 2.1: Behavioral characteristics of young adults who find themselves in financial difficulties

	Passive	*Active*
Problem caused by the individual	*The absent-minded:* Little to none control or overview concerning personal finances	*The shopper:* Short-sighted individuals that constantly overdraw their accounts
Problem caused by external forces	*The victim:* Individuals that feel they have been tricked by the "system", i.e. complicated financial systems	*The juggler:* Operates within the system with relative ease and competence, but will inevitably end up in trouble

(Source: Brusdal & Berg, 2011, p.39)

1. The absent-minded:

These are young adults whose financial problems stem from a lack of knowledge and negligence concerning their personal finances. They have a tendency of not opening their bills, forget or misplace them, and will typically not engage in mental accounting.

2. The shopper:

These individuals will get into financial distress mainly due to a consumption pattern that is not supported by their income. They are impulsive shoppers that are not conscious how excessive spending will affect their personal finances. Future consequences are disregarded, which often results in overdrawn accounts and increasing debt.

3. The victim:

The victims blame their financial problems on complex financial systems and the availability of easy credit. In their opinion, commercial banks and credit companies are too vigilant in giving out credit, and the victim feels "forced" to use the credit cards or credit opportunities they receive.

4. The juggler:

These individuals have extensive knowledge about how financial systems and credit markets function. They use this knowledge to acquire and postpone financial obligations. Often they will juggle several credit loans and credit cards to spread the debt, to avoid facing consequences with any of the companies they are juggling. Also, they tend to pay off outstanding credit with credit from other institutions. Much like Ponzi-financing, this is only sustainable for some time before it all comes to a halt.

These four categories do not consider individual differences and circumstances, but represent stereotypical traits that describe generalized mindsets that recur in the study. They serve to give an overview of how different mindsets may lead to the same consequences. Individuals who have traits falling under each of these categories may affect the outcome of our study.

The objective of our thesis is to test whether it is possible to influence the decision-making process of young adults. In doing this, we are testing if we can reduce propensity to default on payments by using a fairly cheap tool without introducing forced regulations; namely nudging people's behavior. We will now introduce the theoretical perspective for this thesis and go more in depth on the theories that support our approach to this issue.

2.2 Making the right decisions

Making decisions, or rather, making wise decisions is an aspect of life that most individuals are trying to master in the best possible way. We are faced with an array of decisions every day. Some decisions are trivial and simple while others are life-changing and very difficult to make. We make decisions continuously as we go about our day, standardizing some of the decisions to become subconscious choices. But we are sometimes faced with decisions that need to be made only once in a lifetime. Although we have a sense of what it means to make decisions, Jonathan Baron (2008) has defined a decision as follows:

«A decision is a choice of action - of what to do or not to do. Decisions are made to achieve goals, and they are based on beliefs about what actions will achieve the given goals» (Baron, 2008, p.6).

Given that decision-making means choosing actions to achieve our goals, how should individuals go about choosing their actions? To answer this question we introduce theoretical models related to the field of judgment and decision-making.

There are three different categories of models that are used in different fields of study; descriptive models, prescriptive models and normative models. Within judgment and decision-making, descriptive models are models or theories that describe how people think and behave (Baron, 2008). E.g. how most people make a given decision. Prescriptive models are models that explain how individuals ought to think and act (Baron, 2008). E.g. how most people should make a decision compared to how they actually do make the decision. To be able to decide which prescriptive models are most useful we apply a normative model, which is the standard that defines what ways of thinking and choosing are best for achieving our individual goals (Baron, 2008). Normative models evaluate thinking and decision-making in terms of personal goals. In the field of judgment and decision-making the normative model will consist of the policy, or pattern of choices, that will achieve these goals to the greatest extent in the long run (Baron, 2008).

The normative model for decision-making dictates that people should pay their bills on time in order to avoid late-fees and undesirable ripple effects in their personal finances. However, observations show that a significant number of people default on bill payments every month. This can be viewed as individuals departing from the normative model for decision-making, and represents the descriptive model for our case. Prescriptive models describe ways to lessen the gap between the normative and descriptive model; how can we influence people to pay their bills more promptly? One possibility is forced methods; either by introducing pre-paid subscriptions only, automatic draws for people's bank account, or closing the subscription the following day if payment is not made. Here, we are testing a prescriptive model aiming to preserve freedom of choice while giving nudges to influence people's choices, giving them a mild push in the right direction.

Most people strive to reach their goals in life. We wish to have great careers, own property, get married and have children, and live happy and fulfilling lives. Financial difficulties and reduced credit scores may hinder us from achieving such long-term goals. If people made their choices

exclusively according to the normative models for decision-making, they would, in theory, achieve all their long-term goals in life.

2.3 Biases and heuristics

One of the main problems related to human decision-making is that we are not always able to make decisions that will best achieve our goals in the long run. Despite having good intentions, we observe bad decisions being made all around us. People choose to smoke, eat unhealthy foods, drive faster than the speed limit, skip work, or neglect to pay their bills on time. These decisions are made in spite of the obvious fact that the choices will have negative consequences. So why do we make such choices, even though we are at least partially aware of immediate or future consequences?

We can find some answers by looking into previous research conducted in the field of judgment and decision-making. In "Heuristics and Biases" (Kahneman, Gilovich, & Griffin, 2002), the authors discuss how the models and understanding of decision-making have changed over time. For a long period of time the classical model of rational choice was the prevalent view on how individuals make their decisions (Kahneman et al., 2002). This model has been prevailing in the field of economics and has also had considerable influence in behavioral and social sciences (Kahneman et al., 2002). The model of rational choice stipulates that a rational actor, which is represented by any given individual, will choose what option to pursue by assessing the probabilities of different outcomes, judging the utility that can be derived from each possible outcome, and combining these assessments to make their choice. The optimal combination of probability and utility is the option that will be pursued (Kahneman et al., 2002). Calculating these probabilities and multiattribute utilities may be difficult, but the rational choice model still assumes that people are able to do it, and that people do it well. The theory does not claim that mistakes are never made, but the theory is insistent on the fact that mistakes are unsystematic (Kahneman et al., 2002).

With time, this model came under scrutiny. Several researchers made contributions to the judgment and decision-making literature that made the classical model loose much of its foothold. According to Grove (2005), Paul Meehl (1954) found that there were significant discrepancies between clinicians' assessment of their own performance and their actual records of success. These irrational assessments inspired further research on the faulty inferences made

by individuals (Grove, 2005). In 1964, Ward Edwards introduced Bayesian analysis into the field of psychology, which provided a normative standard with which trivial decisions could be compared (Edwards, Lindman, & Savage, 1963). From Edward's own research it became clear that intuitive judgments by individuals did not correspond with the normative standard (Edwards et al., 1963). Herbert Simon also developed significant theories in this field. In his work from 1955, Simon concludes that the "omnipotent" rationality, as described by the rational choice model, was an unrealistic standard for human judgment (Simon, 1955). He proposed the concept of bounded rationality, which acknowledged the processing limitations of the human mind. The concept of bounded rationality states that people make rational choices, but only within the constraints of limited search, knowledge and computing capabilities (Simon, 1955).

Later, Kahneman and Tversky would develop their own unique perspective on bounded rationality, inspired by the examples of biased real-world judgments by the aforementioned authors. In their most influential work from 1982, Kahneman and Tversky explored different heuristics and how these may lead to systematically biased decisions (Kahneman, Tversky, & Slovic, 1982). A heuristic is explained to be a rule of thumb; a simple and more or less efficient rule which people tend to use when forming judgments and making decisions (Kahneman et al., 1982). In most situations people do not have the time, resources or capabilities to make exhausting searches for information. We therefore use heuristics as mental shortcuts, often only focusing on very few aspects of the decision while discarding others (Kahneman et al., 1982). Such heuristics have proven to be very useful in many situations, and the consequences of making decisions based on heuristics are usually trivial. E.g. if you know that one handful of pasta feeds one person, you assume that four handfuls will feed four people. You do not take the time to measure the amount of pasta accurately, nor do you research the nutritional values of pasta to fit a larger group of people. The problem occurs when these heuristics lead to individuals deviating from normative theories, making systematically biased decisions. Biased decisions can be described as departures from the normative rational theory that serves as markers for related heuristics (Kahneman et al., 2002).

2.4 Choice architecture

Decisions are not made in a vacuum. Research performed on biases and heuristics has shown that the context in which individuals make their decisions is significant for the decisions made. Researchers therefore point to context-dependent factors as ways to influence individuals' decision-making process. Marketers frequently use this to influence the purchase decision. How the different choices are presented in the store, what information is conveyed in the advertisement, how the sales process is structured, are all examples of how marketers attempt to change the context to influence individuals' purchasing decisions in a specific direction. While marketers use this knowledge to promote their own interests, some also aim to use this knowledge to benefit society as a whole. Organizing the context in which we make decisions is called "choice architecture", a term coined by Thaler and Sunstein (2008). E.g. when designing employee satisfaction forms for a company, you are a choice architect. When designing the layout of a grocery store, you are a choice architect. When putting forward a legal defense in front of a jury, you are a choice architect. By changing the context or setting in which individuals or groups make their decisions, you may have a significant effect on the decisions that are made. This means that everyone, from parents presenting bedtime options to their children to a government providing various policies to its citizens, influence choices and actions taken as choice architects.

Choice architects have many different ways to influence choices: e.g. by varying the presentation order of alternatives, how to structure the different alternatives, the order of attributes, or changing the default option (Johnson et al., 2012). While it might be tempting to believe that choices can be presented in a "neutral state", the reality is that there is no neutral state. Any way that the choice is presented and conducted might influence the choice of the decision-maker (Johnson et al., 2012). Consider the following: All choices will have a default option; the selected option when the decision-maker chooses not to make a choice between alternatives. A given option is more likely to be chosen if it is set to be the default option (Thaler & Sunstein, 2008). This means that in many cases, a choice architect – being any given individual, can increase the probability of an alternative being chosen simply by making this alternative the default option.

Given that we have the ability to influence choices by changing the environment, there are also possibilities to enhance the quality of the decisions made. In their book, "Nudge", Thaler and

Sunstein (2008) discuss how choice architects should alter the environment in which people make their decisions, in order to nudge people into making decisions that are objectively better for themselves in the long run (Thaler & Sunstein, 2008). If people made better decision for themselves, this may in turn benefit the entire community.

2.5 Nudging – on how to improve decisions

To nudge is formally defined as *«to push mildly or poke gently in the ribs, especially with the elbow – to alert, remind, or mildly warn another»* (Thaler & Sunstein, 2008, p.4). However, Thaler and Sunstein uses the term about facilitating the environment in which people make their decisions so that people can make better choices; giving them a slight "push" in the direction of making a better decision than they otherwise might have. A nudge is any aspect of choice architecture that influences people's behavior in a predictable way without denying any options or significantly changing incentives. To count as a mere nudge, the intervention must be fairly easy and cheap to avoid (Thaler & Sunstein, 2008). If it is not, it simply crosses over to functioning as a forced regulation. Nudges are not orders; while putting fruit items at eye level or closer to the cashier is a nudge, banning junk food is not. A nudge will not influence "Econs"; unbiased individuals who respond to incentives only, but will significantly alter the behavior of "Humans"; biased individuals who respond to incentives as well as nudges (Thaler & Sunstein, 2008). By properly making use of both incentives and nudges, one may be able to help solve many of society's problems while ensuring individuals' freedom of choice.

People might need a nudge in different situations. Usually, people tend to need nudges when faced with decisions that are difficult and rare, for which they do not get prompt feedback, and when they struggle to translate the aspects of the situation into easily understandable terms (Thaler & Sunstein, 2008). Also, we might need nudges in situations where it is difficult to exercise self-control; typically when the consequences of a certain choice are not experienced until some time has passed (Thaler & Sunstein, 2008).

There are six main categories of nudges, defined by Thaler and Sunstein (2008), which are further explained below. Each type of nudge is typically used for different purposes and contexts. The objective of using any kind of nudge is always the same; helping individuals to make better and more informed choices.

iNcentives – making incentives salient if they are hidden. E.g. credit card companies have a tendency to make interest and contractual information difficult to find and difficult to understand for the general population. By properly informing customers about prices and conditions, people would be better equipped to make well-informed choices.

In terms of bill payments, one could introduce concise and detailed information sheets that are sent out to all new customers when applying for subscriptions. By doing this, customers would become aware of the possible consequences of late payment at an early stage, and this would hopefully help reduce late payments.

Understand mapping – helping people to structure complex choices, to enhance their ability to choose the option that will make them better off. An example could be RECAP (Record, Evaluate, and Compare Alternative Prices), which is a program making it possible for consumers to find and compare prices from many different suppliers. This type of nudge is often used by consumer organizations by offering customers help with calculations.

Helping people with mapping could also be used in the context we are currently researching. One could make a RECAP program for the particular industry, where consumers could compare and review all the different service providers to see which company that would best suit their needs. By helping people choose the best suitable provider, one might reduce default payments.

Defaults – people tend to go for the default options, because of inertia, the status quo bias, laziness or other reasons. Choice architects may use this by letting the best outcome in the view of the receiver be the default option, e.g. making the safest pension plan the default in a company.

Making automatic bill payments the default payment option, could help reduce defaults for customers. People tend to go with the default option, and not make active choices. Having more customers using this feature one could reduce the amount of default payments.

Give feedback – letting people know when they are doing well and when they are not, and give warnings when poor decisions are detected in time to change them.

Giving customers positive reinforcement when they pay on time, or warning them when they are about to default payment, may influence the customer to keep paying on time in the future, or giving them the opportunity to correct their mistake if they are about to default. This can be done by emails or text messages, informing the customer that the payment has been received, or

sending out a message some time before the payment is due to remind the customer about a bill that needs to be paid.

Expect error – arranging choices so that it is difficult for people to make mistakes. E.g. the diesel nozzle does not fit gasoline cars, making it impossible for people to fill their vehicles with the wrong fuel. Another example of this type of nudge is ATMs not giving out cash before the card is removed. By having people remove their card first, the probability of leaving ones credit card in the machine is reduced.

If the company expects that a large proportion of their customers will pay the bill late, they could set the due date on the invoice to be some time before the bill is actually due. If the customers view the due date as the 15th, but the company does not consider the payment defaulted before the 20th, they are adjusting for expected errors.

Structure complex choices - finding simplifying strategies that will ease information gathering and choice.

This type of nudge is very similar to *Understanding mapping* as this type of nudge aims to find ways to simplify information search for people to make choices. One example of this type of nudge related to our case is the website www.telepriser.no. By inserting information about typical usage pattern, this site helps navigate the jungle of providers by suggesting the subscriptions that are best suited for the user's needs.

(Thaler & Sunstein, 2008)

In their work, Thaler and Sunstein present many examples of how it is possible to use nudging as psychological tool to influence the decision-making process. These examples are mostly of anecdotal nature, giving brief summaries of previous studies or short stories about human quirks that can be observed in real life. To have any expectations as to what we can find in our study, we need to look further in depth on what research has previously been performed and the results of these studies. By specifically looking at experiments performed in settings similar to ours, we hope to gain a clearer picture of what we can expect in terms of results from performing our study.

2.5.1 Review of relevant experiments

When we started searching for past research on this topic, we found that most articles on influencing the decision-making process and behavior were focused on incentives. More specifically, much of the research focused on monetary incentives and how this may be used to influence decisions. Charness and Gneezy (2009) found that it was possible to significantly influence and enhance individuals' exercise routine by offering financial incentives in a given time period. They also found that positive habits formed under this benefit scheme were continued after the monetary incentive seized to exist. Similarly, Angrist and Lavy (2009) found that it was possible to greatly increase the number of female students that obtained their high school diplomas by offering cash incentives for final graduation. There is an abundance of research that supports a relationship between financial incentives and changes in behavior, and it seems to be a very effective way to perpetuate good habits. However, monetary incentive schemes do not fall under the definition of nudges, as they do not align with the characteristics of a nudge; they are not easy to implement and they are not cheap to avoid. For our review of relevant experiments, we have therefore chosen to focus on research where individuals' decision-making process is influenced by other measures than financial incentives.

An initial problem we were faced with is that "nudging" is a relatively new expression. The term was first coined in 2008 and the following research on nudging has been limited. But, when drawing on the core principle, namely attempting to influence decision-making processes by altering the environment surrounding individuals, there are many helpful studies both prior to and after the term itself was defined. Even though none of the articles we have investigated mention "nudging" specifically and the theoretical perspectives may vary to some extent, the basic principles are quite similar.

To shed light on what results can be expected from our experiment we have chosen to focus on five articles that test the effect of influencing financial choices by the use of non-financial tools. For the purpose of comparison, all the experiments conducted in these five articles use the same dependent variable; willingness to pay. The research is conducted in different contexts and differs in independent variables; what is proposed to influence willingness to pay.

On the next page is a matrix structuring the differences in contexts, independent variables and relevant findings of the five selected studies we found relevant to our experiment.

Table 2.2: Recap of selected relevant experiments

Article	Context	Dependent variable	Independent variable	Findings
Applying behavioral insights to reduce fraud, error and debt (CAB 099-12, 2012)	A division within the UK Cabinet Office aimed to influence late tax payers to pay their taxes.	Willingness to pay.	Nudging based on social norms.	Nudging based on social norms increased tax compliance by 15 %.
Globalization issues and consumers' purchase decisions for food products (Disdier & Marette, 2013).	A laboratory experiment to investigate how globalization issues influence willingness to pay for food products.	Willingness to pay.	Positive/negative information about MNC's.	Negative information about a MNC's operation significantly decreased willingness to pay for food products – and vice versa.
Effects of social responsibility labeling and brand on willingness to pay for apparel (Hustvedt & Bernard, 2010).	A study examining changes in university students' willingness to pay for apparel as labor-related information is added to products.	Willingness to pay.	Presence and nature of labor-related information.	Consumers were willing to pay more for apparel containing information about labor-related attributes of the apparel.
What's in a name? (Menegaki, Mellon, Vrentzou, Koumakis, & Tsagarakis, 2009)	A study examining how the agricultural sector's and consumers' willingness to pay for putrefied water changes with descriptive terms concerning the product.	Willingness to pay.	Descriptive terms / framing of alternatives.	Farmers' and consumer's willingness to use and pay for putrefied water increased when it was called "recycled water" instead of "treated wastewater".
Does attribute framing in discrete choice experiments influence willingness to pay? (Kirsten & Glenn, 2009).	Experiment conducted to examine whether attribute framing influences willingness to pay for different cancer screening alternatives.	Willingness to pay.	Framing the presentation of alternatives.	By framing outcomes for a given cancer screening method in a positive manner, the researchers increased the likelihood of that treatment being chosen – and vice versa.

We have focused our review of past experiments on studies that aimed to influence or change attitudes and/or behavior in the decision-making process concerning financial choices. In the matrix we have listed willingness to pay as the dependent variable. Willingness to pay is an expression that can be used to describe different concepts. The article about tax compliance in the UK looked at behavioral changes, and individuals' willingness to actually pay the taxes they owe. The article about globalization issues related to food products, and the article about social responsibility labeling of apparel, studied how it is possible influence attitudes and opinions related to how much people are willing to pay for a product. The article about treated wastewater investigated if it is possible to influence consumers' attitudes and opinions related to their propensity to use or pay for a certain product. Lastly, the article about preferences related to screening methods for cancer examined if the framing of risk probabilities may influence preferences and how much people are willing to pay for different services. Even though these articles are looked at different aspects of willingness to pay, they all attempted to uncover whether it is possible to influence the decision-making process, both in terms of attitudes and actual behavior, by altering the environment surrounding the decision maker. The goal of our thesis is to see whether we can use nudging as psychological tool to influence people's willingness to pay their cell phone bills, and subsequently reduce default rates.

After reviewing these articles we find that the experiments conducted were all successful and had significant results. Even though the research is conducted in different contexts, using different independent variables, they all conclude that it is in fact possible to influence consumers' willingness to pay by implementing non-financial measures. From these five studies we can conclude that people are not as in control of their own decision-making as they might think they are. The surrounding environment, combined with heuristics and biases, will in most cases influence people to a greater extent than one might think. In all the presented articles the authors show that a wide array of factors may influence the decision-making process and consumers´ willingness to pay.

«Applying behavioral insights to reduce fraud, error and debt»

The basis for this experiment was increasing problems with tax fraud and error in declarations and debt in the UK. Fraud, error and overdue debt accounts for net losses that amounts to approximately 40 billion GBP each year (CAB 099-12, 2012). According to the Behavioral Insights Team (BIT), traditional means to solve these problems assume that individuals make decisions according to the classic model of rational choice (CAB 099-12, 2012). Rational actors will weigh the costs and benefits of committing fraud, default on debt and hand in declarations with errors, and refrain from doing so it the costs outweigh the benefits (CAB 099-12, 2012). BIT aimed to examine if there was any way to utilize knowledge from behavioral sciences to help reduce the billions in lost tax revenue. The researchers found the best results when using social norms to influence late taxpayers to pay. Letters were sent to 140.000 late taxpayers, where some received control letters (a reminder to pay their outstanding taxes) while others got letters containing a reminder combined with statements developed according to social norms (CAB 099-12, 2012). These social norms were statements like "9 out of 10 people in Britain pay their taxes on time", or that "most of the neighbors have already paid their taxes" (CAB 099-12, 2012). Referring to a social norm in a particular area resulted in a 15% increase of payments compared to the control letters (CAB 099-12, 2012). This result was significant and could generate 30 million GBP in extra tax revenue each year (CAB 099-12, 2012).

"Does attribute framing in discrete choice experiments influence willingness to pay?

The authors of this article stipulated that when attributes of a certain choice is coupled with different risks, an individual is required to process information about probabilities and utility values within the bounds or rationality (Kirsten & Glenn, 2009). The authors aimed to test whether the way these risk probabilities are framed can influence an individual's decision-making process. To test this, an experiment was designed to test if the framing of risk attributes related to different screening alternatives for colorectal cancer could influence patients' willingness to pay for a given screening alternative (Kirsten & Glenn, 2009). This was tested by mailing surveys to 1920 test subjects enrolled in the Central Coast Rotary Bowelscan Program in NSW, Australia. The participants were randomly assigned one of four alternative types of framed information, which framed risk attributes differently; either positively or negatively (Kirsten & Glenn, 2009). One of the alternatives focused on if the accuracy of a screening method for cancer

that was framed as either "How many cancers the test finds" or "How many cancers the test will misses". The probabilities were the same, but one was expressed with a positive outcome and the other with a negative outcome. Another alternative framed the risk associated with a screening method as either "The number of people who are correctly reassured that they do *not* have cancer" and "The number of people who have unnecessary colonoscopies" (Kirsten & Glenn, 2009).

The study supports the authors' assumption that framing of outcomes and risk probabilities may have a significant influence on an individual's willingness to pay for a given screening option. The authors found that individuals had increased willingness to pay for a given screening method if it was framed as cancers found, rather than cancers missed (Kirsten & Glenn, 2009). The same was the case for the other aforementioned alternative; test subjects had significantly increased willingness to pay if a screening method was framed as how often the screening method could rule out cancer, rather than framing it as how often people would get an unnecessary colonoscopy (Kirsten & Glenn, 2009).

«Globalization issues and consumers´ purchase decisions for food products: evidence from a lab experiment»

In this study, the authors examined whether factors related to globalization could affect willingness to pay. More specifically, whether information concerning a MNC's operations could affect willingness to pay (Disdier & Marette, 2013). A laboratory experiment was conducted, using a sample of 101 men and women aged 20-72 (Disdier & Marette, 2013). Each participant was to indicate what price they would pay for a glass of pickles from a traditional and well-known French brand, Maille. They were informed about a reference price; the average retail price for such a jar of pickles (Disdier & Marette, 2013). After being given a reverence price, the test subjects were asked to indicate what price they would be willing to pay for this jar of pickles. After having stated their price, they were given more information about Maille's operations. First, they were told that Maille, after being bought up by Unilever in 2000, had outsourced the supply of pickles from France to India. Next, they were told that Unilever had closed down processing facilities in France due to outsourcing (Disdier & Marette, 2013). The third round of information conveyed positive information about new products and services from Maille. And lastly, the participants were informed about new investments in and by Maille. After each round

of new information, the participants were to indicate the maximum price they were willing to pay for a jar of Maille pickles (Disdier & Marette, 2013).

The study found that both positive and negative information about the firm's operations could have significant effect on a subject's willingness to pay (Disdier & Marette, 2013).Their findings regarding the difference in average willingness to pay (WTP) was found to be:

1. After being informed about retail price: €2,75 per jar
2. After being informed about outsourcing supply: €2,21 per jar (-19,63% from retail price)
3. After being informed about the closing of facilities: €2,06 per jar (-25,1% from retail price)
4. After being informed about new products and services: €2,31 per jar (+12% from the former price)
5. After being informed about future investment activities: 2,68 per jar (+31% from the lowest WTP)

    (Disdier & Marette, 2013)

The researchers noticed that negative information about a company's operations would negatively affect WTP among consumers, but this effect was offset by positive information. After the subjects were told two negatives, and two positives, the WTP (€2,68) was very similar to the initial WTP (€2,75). This showed how consumer WTP may be significantly affected by information provided, and opens for the possibility to influence the decision-making process for consumers.

«What´s in a name: framing treated wastewater as recycled water increases willingness to use and willingness to pay»

As the title states, this article looked into the decision-making process of individuals to examine whether individuals were more prone to accept a given product if the name/terminology was changed. The product in question was treated wastewater, i.e. treated sewage water (Menegaki et al., 2009). The background for performing this study was the increasing problems with water shortage in the agricultural sector in Crete, Greece. Treated wastewater water could help increase the water supply used for irrigation, yet most people had been reluctant or unwilling to adopt this source of water (Menegaki et al., 2009). The main reason for farmers' reluctance to use treated

wastewater for irrigation, and consumers' unwillingness to pay for foods grown with the use of this type of water, was that they perceived it as being dirty or unhygienic (Menegaki et al., 2009). People seemed to have a difficult time accepting the use of water that has previously contained excrement and fecal matter, regardless of the clear evidence provided of the water's purity (Menegaki et al., 2009). By reframing this product as "recycled", the authors wanted to see if this could increase willingness to pay or willingness to accept. The experiment was conducted in Crete, Greece, using 1004 test subjects (Menegaki et al., 2009). This sample included both consumers and farmers. The subjects were asked to answer questions related to their willingness to use and willingness to pay for the product. Half of the respondents were given questions where the product was labeled as treated wastewater, while the other half were given questions were the product was labeled as recycled water (Menegaki et al., 2009). The results showed a significant difference in response caused by this difference in labeling. For each question asked, there were four alternative answers: definitely no, probably no, probably yes, and definitely yes. The findings can be summarized as follows:

1. On average, 37% answered definitely no when asked if they would buy or use treated wastewater. Only 14% gave a definite no when it was labeled as recycled water.
2. On average, 28% answered probably not when asked if they would buy or use treated wastewater. Only 14% gave a probable no when it was labeled as recycled water.
3. On average, 21% answered probably yes when asked if they would buy or use treated wastewater. But 34% gave a probable yes when it was labeled as recycled water.
4. On average, 12% answered definitely yes when asked if they would buy or use treated wastewater. A staggering 30% gave a definite yes when it was labeled as recycled water.
   (Menegaki et al., 2009)

Based on these findings the authors argued that the way in which a product is portrayed or framed towards potential users and consumers could have a significant effect on an individual's willingness to use or pay for a certain product. This seems counter intuitive, because the people in the experiment were fully aware that treated wastewater and recycled water was exactly the same product. According to the rational choice theory, individuals should not be affected by *how* a product is presented, only the intrinsic value of the product. This seemed to not be the case in

this experiment; simply altering the name of the product made possible to affect farmers' willingness to use and consumers' willingness to pay for the specific product.

«Effects of social responsibility labeling and brand on willingness to pay for apparel»

This article examined whether or not it was possible to increase consumers' willingness to pay for clothing items by conveying labor-related information on apparel. The study was performed at Southern University, US, using 121 university students as subjects (Hustvedt & Bernard, 2010). The background for this study was the fact that companies are forced to make their business practices more in line with ethical codes due to increasing consumer awareness and the possibility to lose customers due to unethical practices (Hustvedt & Bernard, 2010). Based on this emerging trend, the authors aimed to examine whether this attitude towards ethical awareness also could affect willingness to pay (Hustvedt & Bernard, 2010).

Hustvedt and Bernard tested this by staging 14 experimental auction sessions. Students from a variety of departments and collages participated in the auctions. The students were given $40 for participating, and could either use this money to place bids in the auction or take home as cash. In a series of four steps, the students were to bid on t-shirts with no information or branding, then on t-shirts with labor information and lastly to bid on t-shirts with labor information and a brand name (Hustvedt & Bernard, 2010). The authors found that the students' willingness to pay was significantly increased when there was given information about labor conditions on the apparel in the auctions (Hustvedt & Bernard, 2010).

The aim of this review has been to look back on past research to see if we could uncover trends or results that are relevant for our thesis. We wanted to see whether past research could shed light on what results could be expected from experiment we will be conducting.

We have explained choice architecture as being any changes of the environment in which people make their decisions (Thaler & Sunstein, 2008). From studying past experiments it seems that the researchers are consistently able to alter people's attitudes, decision-making process, and ultimately behavior, by applying some sort of external stimuli. As we will be testing the effect of external stimuli on a decision-making process similar to the ones addressed in the presented research, we are hopeful that we could find similar results in our study.

As mentioned in the introduction to the review of past research, the authors of these articles did not use the term "nudging" specifically, but they did, however, use the same principles that were put forth by Thaler and Sunstein (2008); they found unrestrictive and cheap methods to change the environment in which the test subjects made their decisions. This influence, or nudging of behavior, was shown to alter attitudes and behavior in both laboratory experiments and field experiments. We are under the impression that the findings could easily be transferable to our experiment.

With our thesis and experiment, we aim to contribute to this growing field of research. We want to test nudging in a context where it has not to our knowledge been tested before; influencing willingness to pay, in the form of actual payment behavior, were the test subjects answer to a privately owned company, not the state. Our goal is to lower the number of young adults that default on their bills, serving a sociopolitical purpose as well as the purpose of bettering personal finances for individuals.

2.5.2 The nudges used in our study

To be able to develop nudges that are appropriate to use for our study, we have looked further into the heuristics and biases that we believe may influence payment behavior. We found three heuristics and biases that we believe give insights to how we should structure our nudges in text format, to be empirically tested in our specific context. The idea was to send out nudges in text format with different content to independent groups of test subjects, to investigate whether the information conveyed could influence the decision-making process and produce more people to

pay their bills within due. Our nudges do not specifically fit either of the nudge categories presented by Thaler and Sunstein, but they are developed in the spirit of *INcentives*, **Give** **feedback** and **Expect error**.

2.5.2.1 Nudging using a common reminder

We wish to test an additional nudge to the nudges specifically developed according to theory on heuristics and biases; a common reminder. By reminding people about the bill, just before it is due, we hope that less people will default on their payment purely because they forget the upcoming due date. So this nudge, or message, will not contain any additional information developed according to theories on heuristics and biases. Research has shown that that reminding individuals of upcoming tasks or events significantly increases the probability of it being followed through. Calzolari and Nardotto (2011) found that it was possible to increase attendance at gymnasiums by up to 25 % by sending out weekly e-mails reminding individuals of the possibility to attend the gym. Karlan, McConnell, Mullainathan, and Zinman (2010) found that it was possible to increase people's savings for future expenditures by simply reminding them that it would be likely that they would experience higher expenditures in the future than they had at the time.

We believe that a reminder will have a positive effect on reducing default rates. It will be especially interesting to observe if there is any difference in effect from sending out a common reminder compared to nudges developed according to heuristics and biases. If the messages we distribute produce different results, this would indicate the importance of how the nudges are coded and formulated in text format.

We set up the hypotheses to test the relationship between nudging using a common reminder and the default rate as follows:

H_0: There is no difference in default rates between the group that receives the common reminder and the control group

H_A: The default rate will be lower in the group that receives common reminder compared to the control group

The following nudges will contain added information to the simple reminder. We expect that by adding information based on our knowledge of heuristics and biases, these nudges, or messages, will have a greater effect on default rates than the reminder alone. In addition to reminding the customer of the upcoming bill, we add incentives to pay the bill within due.

2.5.2.2 Nudging based on the anchoring-effect

When people are faced decisions that require that they exercise judgment, one common heuristic is to use a given number or sum as an anchor, and adjust their judgments according to this number (Kahneman et al., 1982). If you are selling your bike and you know that a friend has sold a similar bike for $500, you might decline an offer lower than $500 because you judge it to be too low compared to the anchor sum. This is effect is also prevalent regardless of the relevance of the number used as the "anchor". For example, one experiment was performed asking people to estimate the percentage of African countries in the UN. Before giving their answer, a wheel of fortune containing numbers between 0-100 was spun in the test subjects' presence. It became clear that the number that appeared on the wheel actually influenced their estimates on the percentage of African nations in the UN. Participants with higher number on the wheel consistently gave estimates that were significantly higher than participants who got low numbers on the wheel (Kahneman et al., 1982). When people anchor in numbers that are not arbitrary, i.e. that are in fact relevant for the judgement at hand, this effect has shown to be even larger (Kahneman et al., 1982). If the anchor is perceived to be relevant to our issue, the anchor may significantly affect our final judgement.

Based on this theory, we wanted to structure a nudge that gave young cell phone customers a reference number, in our case the number zero. The message we wanted to send out conveys information that the customers will only pay for the actual use of services - the product itself, and will face no extra fees or costs given that they pay on time. By anchoring the customers in the number zero, i.e. only paying for use, we hope to provide a reference number that will affect the judgement on whether or not they pay their bill within the due date. We hope this reference number will serve to remind customers that it is only if they pay on time that there are no extra costs added to their final bill.

We set up the hypotheses to test the relationship between nudging based on the anchoring-effect and the default rate as follows:

H_0: There is no difference in default rates between the group that receives the nudge based on the anchoring effect and the control group

H_A: The default rate will be lower in the group that receives nudge based on the anchoring effect compared to the control group

2.5.2.3 Nudging based on loss aversion

In 1979, Kahneman and Tversky introduced a descriptive theory of decision utility called "prospect theory". This theory gives insight as to why people are loss averse. The theory explains that the negative feelings an individual experience from a given loss is greater than the positive feelings the same person would experience from an equal gain. Most people would perceive the loss from losing NOK 100 as much greater than the gain from receiving NOK 100. This is why many people would not accept a bet with an equal chance of losing or winning the same amount of money, even if the expected value of the bet is zero. Some studies suggest that losses are as high as twice as psychologically powerful compared to equal gains.

This theory inspired us to develop a nudge based on people loss aversion. We formulated a nudge were the customer is informed of the minimum charge that is added from defaulting on payment, which is NOK 63. NOK 63 is the maximum rate for first-time late-payment fee in Norway, and is used as the standard fee in most industries. This sum of money is fairly small and manageable for most people when isolated, but prospect theory stipulates that people are loss averse even when it comes to smaller sums (Kahneman & Tversky, 1979). When made aware of the potential loss people might weigh this against potential gains they might achieve from not paying their bill on time such as spending their money elsewhere. We expect that people will perceive the potential extra cost as worse than any possible gain they may have short-term by defaulting payment. They will have to pay the amount owed at some point; postponing payments will only result in greater costs.

We set up the hypotheses to test the relationship between nudging based on loss aversion and the default rate as follows:

H_0: There is no difference in default rates between the group that receives the nudge based on loss aversion and the control group

H_A: The default rate will be lower in the group that receives nudge based on loss aversion compared to the control group

2.5.2.4 Nudging based on social norms

Thaler and Sunstein (2008) explain that humans, unlike econs, are biased and rather easily influenced by the statements and actions of others. Teenage girls are more likely to become pregnant when their friends are having children. This is one of the reasons why one can observe "trends" of teenage-pregnancy in certain communities. One can observe the same tendency for obesity; if your friends or someone in your family gain weight, it is very likely that you will also put on some extra weight. One cannot not exclude other influencing factors that helps explain such trends, but it is a fact that humans are heavily influenced by the people around us (Thaler & Sunstein, 2008). Another interesting example can be found in an experiment conducted in 2011 that looked at the relationship between social norms and energy conservation. In this experiment people received energy consumption reports showing the energy consumption of the recipient, and also the energy consumption of the neighbouring residents (Allcott, 2011). When people were able to track their energy consumption compared to their neighbours, people started to change their consumption patterns. For example, the bracket of test subject that originally had the highest energy consumption, reduced their consumption by 6,3 % on average. Even the test subjects that had the lowest energy consumption before the treatment lowered their energy consumption, even though they could have increased their consumption significantly and still used less than average.

Based on such examples, Thaler and Sunstein (2008) stipulate that if a choice architect wants to shift the behavior of individuals by the use of nudging, it might be as simple as informing people of what others are doing. People tend to look to the actions of others and decide that the action taken by the majority is an attractive choice; it is perceived as the correct way to go. People also have a wish to be associated with or belong to certain social groups, and therefore adhere to the

ruling social norms. Also, it is an easy choice; you don't have to think about or evaluate your options. The risk of this being the wrong choice often has little or no consequences as it is socially accepted.

In the spirit of nudging based on social norms, we developed a nudge that is similar to the nudges used in Allcott's study from 2011, and the experiment performed on tax compliance in Britain in 2012. We wanted to distribute a text message conveying information about how the majority of the cell phone-customers behave; that most people pay on time. Our goal is to investigate whether this information may influence the late-payers and ultimately reduce defaults.

We set up the hypotheses to test the relationship between nudging based on social norms and the default rate as follows:

H_0: There is no difference in default rates between the group that receives the nudge based on social norms and the control group

H_A: The default rates will be lower in the group that receives nudge based social norms compared to the control group

Our thesis focuses on a demographic that has shown to have significant difficulties with making payments on time, namely young consumers. By performing our experiment the aim is to examine whether there is a way to nudge the consumers into paying their bill on time. By altering the environment in which they make the decision to pay on time or not, we hope that these nudges will have a significant effect. The choice of paying the bill on time or not is neither difficult nor rare, but it might be hard to get timely feedback and people might struggle to translate the aspects and severity of the situation into something understandable. The choice of paying on time might also be a situation in which it is difficult to exercise self-control. One might want to forget about the bill and go to the cinema instead. Or maybe the bank account is already empty due to previous purchases? We are trying to give people feedback, or incentives to pay their bills on time. We believe that this is a relevant and sound context to test different nudges, and hopefully find a way to give younger consumers the nudge they need to get their finances in order.

3. Methods

In this section, we move on to the underlying methods for implementing our experiment.

3.1 The research process

> *«Facts do not simply lie around waiting to be picked up. Facts must be carved out of the continuous web of ongoing reality, must be observed within a specified frame of reference, must be measured with precision, must be observed where they can be related to other relevant facts. All of this involves methods»* (Rose and Peterson, cited in Ghauri & Grønhaug, 2010, p.3)

Drawing on the above statement, Ghauri and Grønhaug (2010) describe research as a systematic quest to find truth or answers, altering or extending existing knowledge. All research stems from the existence of something we want to gain more knowledge about. Curiosity could therefore be viewed as the driving force underlying all research and investigation.

Zikmund, Babin, Carr, and Griffin (2010) describe two types of research based on the purpose of the study; applied business research and basic business research. Applied business research is used to address a specific business decision for a particular firm, while basic business research is used as an attempt to alter or expand existing theory or phenomena, without aiming to solve a specific problem (Zikmund et al., 2010). The two are not entirely separable, however, as basic research is often the basis for later applied research (Zikmund et al., 2010). One common denominator for both types is the systematic use of scientific method. This is when we use knowledge and evidence to draw objective conclusions (Zikmund et al., 2010). For this thesis, we are performing basic business research as the basis for later applied research internally in the collaborating firm.

Figure 3.1: A summary of the scientific method
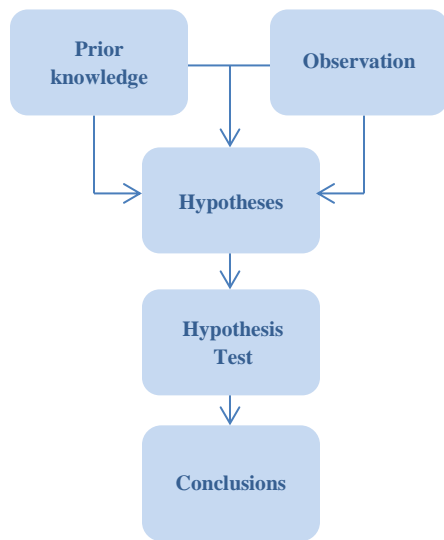
(Source: Zikmund et al., 2010, p.7)

The research process consists of a series of activities undertaken over time. Figure 3.1 outlines the basic steps of the process. There are several ways to develop ideas for further research, which will depend on what we already know and/or what we can observe in real life. We reach the hypothesis stage when we have transformed these ideas into researchable terms. The next step is to test hypothesis against facts or data, resulting in either verifying or rejecting the hypothesis. From this we extract new knowledge in the form of conclusions (Zikmund et al., 2010).

All research will follow a set of phases. In each of these phases choices are made that will affect the following phases and the final validity and reliability of the results. In this section, we will discuss the initial phases of the research process, while leaving the analysis and conclusions to following chapters.

3.2 Problem definition

A concept describes a phenomenon and represents the building blocks of theory; it is an abstraction of reality that is the basic unit for theory development (Zikmund et al., 2010). We will be looking at the concept of "nudging" attitudes and behavior. To clarify our concept we have defined a nudge to be *«anything that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives»* (Thaler & Sunstein, 2008, p.6). The treatments used in the experiment do not limit any options for the customer and can easily be avoided. This is in accordance with the theoretical foundation for this thesis.

The content of a problem definition can be divided into three parts (Jacobsen, 2005). First, we need to define the units we wish to examine. We must then define how to measure the phenomena we wish to investigate. This is done by the use of variables. We expect the units to have different values of the variables. Lastly, we must decide the context of the study. That is, the specific setting and circumstances of the study (Jacobsen, 2005). In our case we will be looking at the current cell phone customers of One Call, aged 18-30, measuring the effect of nudging on payment rates collected from January through March of 2013.

Figure 3.2: The content of a problem definition
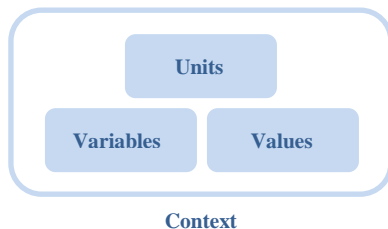
Based on the above definition, we have defined our problem as follows:

We wish to describe if and, subsequently, how the use of nudging may affect the propensity to default on bill payments, i.e. the customers' choice to pay bills within due date. By looking at One Call's customer base aged 18-30, in January through March of 2013, we will test whether nudging may reduce default rates accounted for by this age group.

3.2.1 Hypotheses

A hypothesis is according to Zikmund et al. (2010) defined as a statement that expresses the relationship between two variables, which can be tested empirically. Based on the our research question, we formulate this problem definition into researchable hypotheses to be empirically tested;

H_0: No difference in default rates in treatment group measured against the control group

for m = F, M

H_0: $\mu_C = \mu_i$ for m = F, M

and

H_A: Default rates are lower in treatment group measured against the control group

for m = F, M

H_A: $\mu_C > \mu_i$ for m = F, M

i = {1,4} : treatment groups = 1: reminder, 2: anchoring, 3: loss aversion and 4: social norms.

C = control group

m = month = F: February, M: March

We will measure and test all groups separately against the control group of the corresponding month. The hypotheses will be the same, using the different denotes stipulated above. We thereby test the effect of the different types of messages in the month of treatment, as well as investigating any longer term effects in the following month.

To test the hypotheses, we need to set up a conceptual framework for the study. Any conceptual framework contains a dependent variable and at least one independent variable (Zikmund et al., 2010). The dependent variable represents the phenomenon we want to study, that is to be predicted or explained, while the independent variable represents the factor assumed to predict or explain the dependent variable. Control variables are used in empirical research to reduce the risk of attributing explanatory properties to independent variables that in fact are not responsible for the variation in the dependent variable (Zikmund et al., 2010). The control variable is used to test the possibility that an empirical observed relation between an independent variable and a dependent variable is spurious. A spurious relation is a relation that can be explained by other variables (Andersen, 2012).

We have four independent treatments, being our nudges. These are expected to influence the default rate for each group. We introduce a control group to control for any spurious relationships (e.g. history or seasonal effects). This is a group that will not receive any treatment, and we therefore hope that this group will capture any variation in the dependent variable that cannot be explained by our nudges. We expect that the variation found in the default rates from each group can be explained by our independent variable.

3.3 Unit selection

When we select our units for analysis, we can either choose to gather information from each member of the population or to select a representative sample and generalize the findings to the larger population (Ghauri & Grønhaug, 2010). In the context of our thesis, based on studies and statistics as presented in the previous chapter, and on request from our collaborating firm, the units of interest are limited to cellphone customers in the age 18-30 in Norway. All these customers make up our theoretical population. We do not have the appropriate resources to investigate all the units of the theoretical population. Our actual population is therefore limited to One Call's customers in this age group. We opted to investigate a sample and generalize the findings to count for the total population.

Due to costs and possible ripple effects throughout the organization, One Call made 2500 units available for us to investigate. These were divided into five groups of 500 units, four of which received different text messages, and one was held constant as a control group to eliminate effects of unknown factors. The 2500 were selected from one of three batches of their customers, according to due date of payment. The top 2500 on this list were selected. None of these customers were reported to have changed cell phone provider during the experiment, making our sample size constant throughout the experiment. The list of customers was not structured according to any parameters (Per Ola Stålberg, Customer Developer at One Call, e-mail received 27.02.2013). Since the subjects in our sample have not been selected according to any criteria other than the age demographic, we can assume that we have a randomized sample which is representative for the theoretical population.

3.4 Research design

The research design is the overall plan of the methods and procedures for collecting and analysis of data. The research design will therefore provide a plan of action for the research process. As

several research designs may accomplish the same objectives of a study, the researcher chooses design type according to the appropriateness of each type (Zikmund et al., 2010). This will in turn depend on the objectives of the study, time constraints, the available data sources, the urgency of results, the cost of obtaining data, and assessment of validity and reliability (Zikmund et al., 2010). It is argued that there is no single best research design. However, a specific design type will be more appropriate to use in certain situations than in others.

The purpose of this study is to investigate whether one can find significant effects from nudging customers on their propensity to default bill payments. Therefore it is appropriate to use causal research design. The causal research design is used when we want to explain how events are related; it seeks to find cause-effect relationships (Jacobsen, 2005). In science, this is defined as following: «If X occurs, Y will always happen» (Jacobsen, 2005). Social science has modified this relationship, stating that if A occurs, there is an x % probability that B should occur (Jacobsen, 2005). This requirement is less strict, indicating that a certain degree of regularity rather than absolute correlation is viewed as satisfying. When using this requirement for causality, it is common to operate with three conditions that a study must satisfy in order to make any statements about the causality of a phenomenon (Jacobsen, 2005):

1. Concomitant variation:

There must be a correlation between what we assume is the cause and what we assume is the effect, meaning that the two events vary systematically.

2. Temporal sequence:

Cause must precede effect in time. That is, there must be temporal proximity between cause and effect.

3. Nonspurious association:

The effect must be truly be caused by what we assume is the cause, rather than being due to some other variable.

In addition to these three requirements to judge the causality of a phenomenon, most authors on scientific methods discuss the support of existing theory and research as being important, as well

as other relevant factors and specific conditions that may affect the study. This must also be taken into consideration when assessing the causality of a study.

We wish to use an experiment as our source of data collection. An experiment is *«a carefully controlled study in which the researcher manipulates a proposed cause and observes any corresponding change in the proposed effect»* (Zikmund et al., 2010, p.59). Instead of trying to discover all factors that affect propensity to default payments, we have based our design on the idea that we wish to eliminate all other factors, both known and unknown – only focusing on the effect of nudging. This is the underlying idea of performing experiments (Jacobsen, 2005).

Figure 3.3: Overview of factors that affect propensity to default payments
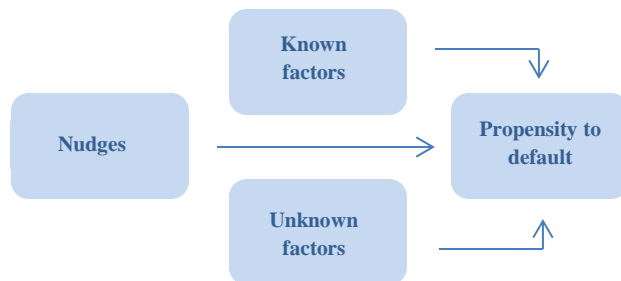
(Based on Jacobsen, 2005, p.111)

The causal ideal is the classic experiment (Zikmund et al., 2010). This design contains four key elements (Jacobsen, 2005):

1. Comparison

This means that we compare changes in the group that has been subjected to the experimental variable with the changes that have occurred in a control group that has not been subjected to the same experiment.

2. Randomization or random selection of units in the compared groups

By randomly drawing the groups to be compared, we hedge against systematic differences between the groups. Thereby we assume that the groups can be directly compared.

3. Time laps data

This means that we register the conditions before the experiment is performed and a similar (preferably identical) registration of the conditions some time after the experiment was performed.

4. Active manipulation

This means that the researcher deliberately manipulates the relationship that is believed to be a possible cause. This relationship is only manipulated within the experimental group, and not in the control group.

Our experiment contains all these elements, as will be further elaborated on in the continuation of this section.

Our experiment can be portrayed as follows:

Figure 3.4: Overview of the experiment
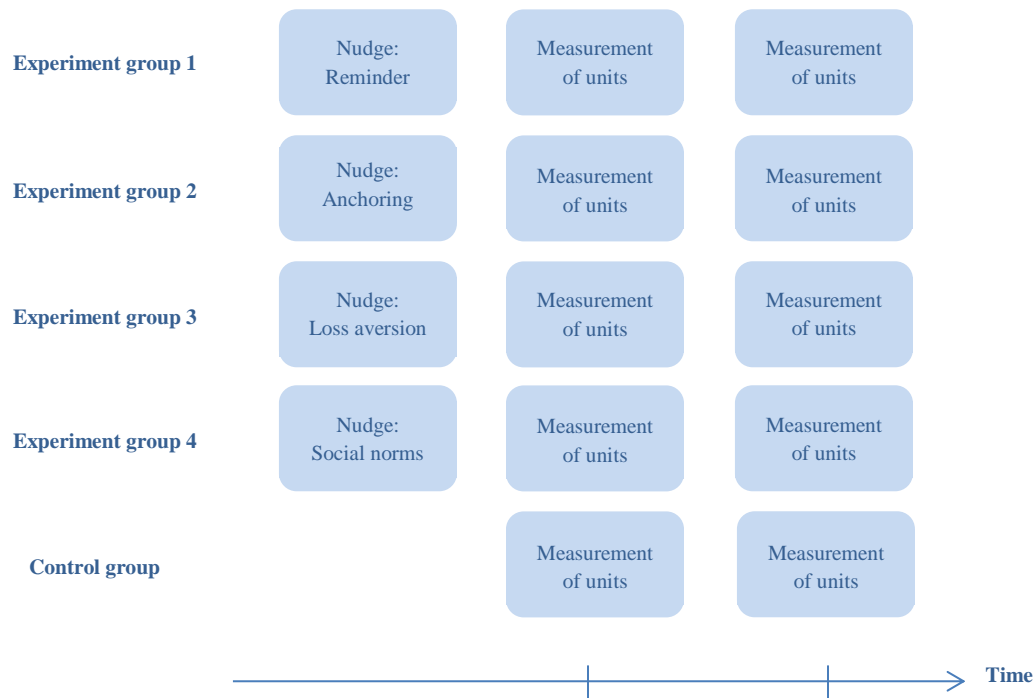
(Based on Jacobsen, 2005, p.113)

In addition to the two times for measurement as shown in the figure we have also estimated, based on total default rates, the expected default rates for each group in January. This represents a measurement for comparison.

The effect of each nudge is calculated as follows:

Experiment Group $_i$ – Control Group = the effect of nudging   for i = {1,4} and m = F,M

The model and calculations show that we have satisfied all three conditions for causality.

3.5 Method for data collection

We have chosen a quantitative research design, as the focus is on the objective facts of a particular phenomenon and is intended to test, verify and extend existing knowledge by testing hypotheses (Zikmund et al., 2010). Quantitative method provides and analyses empirical data in the form of numbers and statistics, and is mostly used when the study aims to describe the extent or frequency of a phenomenon (Zikmund et al., 2010). The main advantage of using this approach is that the information is standardized and is fairly easy to process and analyze by the use of computer programs. Further, the quantitative approach to data collection is usually less costly than qualitative method, which means that we can get many respondents; thereby being able to get a representative sample of the population. From this, we are better able to generalize the findings to count for a larger population (Zikmund et al., 2010). This approach also limits the study by setting a clear finish line in regards to data collection. A disadvantage of using this approach is that we risk getting a superficial study, overlooking underlying factors of *why* we get the specific data. As we will not interview or assess anybody in our sample, it will be difficult to make any assessment of why we see changes or not. Also, we will not register data at an individual level, so we cannot be track one individual's response over time; we are only looking at default rates on an aggregated group level. Another disadvantage is that the researcher has already defined the relevant topics ahead of executing the study, not allowing for any additional factors that might be relevant for the study (Zikmund et al., 2010).

By performing the experiment, we will be gathering primary data. The main advantage by getting primary data is that the data is collected for this particular study, eliminating "noise" from sources having different objectives than ours. The main disadvantage is that it may be difficult to access; collecting primary data can be time consuming and increase costs. The researcher is also

depending on the willingness and ability of respondents to answer truthfully and actually reading the messages or surveys you send out (Zikmund et al., 2010). As we are sending our nudges via text message, we are convinced that we will eliminate most issues related to non-response. In our first meeting with One Call we were informed that in their experience, few of their customers read e-mails or attachments sent by the company. By using such methods, we would face problems with estimating how many actually read the messages. Text messages on cell phones, however, are difficult to ignore as the message pops up on the screen without even opening the message in full. By sending out our text messages, we expect the non-response rate in our sample to be close to zero.

3.6 Procedure

When starting this project we first focused our efforts on finding a company to collaborate with. Without having a collaborating company, and access to its customer base, this project would not be possible to execute. So the first thing we did was to properly formulate what we aimed to do in understandable terms for a firm that might not be familiar with this field of research and our methods for performing such an experiment. We also made a proposal for a progress schedule; how we wanted to proceed and what we needed from the firm at each stage of the process to finish on time. In order to secure that the firm would also get an understanding of the actual treatment, we also formed a suggestion for how we wanted to formulate the messages. These three information sheets made up our "pitch" towards the company: the initial information to send out in hopes of sparking an interest.

We targeted the executives of customer-relations and communication in the companies we approached. Our pitch was sent out to several companies in the cellular industry, based on e-mail addresses either found online or given to us by customer service. The primary reason for looking at Norwegian cellular industry is that it generally has significant default rates with much room for improvement at an individual level. A significant portion of the defaulted bills is accounted for by young adults under 30 years of age. Also, due to time constraints of our project, this is an industry where it would be possible to conduct the designed experiment with rapid and possibly significant results. We got positive feedback from two companies, and after a few rounds of e-mailing back and forward we decided that One Call was the firm we wanted to proceed with.

They invited us to come to their headquarters in Oslo to meet and further discuss the implementation.

Before we met with One Call we conducted a small pre-test in the University cafeteria to verify that we were on to something. This was important both to verify that the coding of the messages was perceived as intended and that there was potential for the expected response. We handed out the different versions of the messages, as proposed in the initial contact towards the company, to 25 randomly chosen students in the cafeteria. After giving them the time to read through the messages, we asked them to give their comments and thoughts. Most of them thought that all the nudges would have an impact on their payment behavior, given that they did not pay their bills promptly. However, most of them stated that this was not an issue for them as they did in fact pay their bills on time. The respondents generally chose nudging based on social norms as the most likely to give the desired effect, so long as it was tastefully formulated not to offend the customer. The common reminder was perceived to be the second most effective message to alter behavior. There were also some respondents that did not believe these messages would have any effect at all. We presented the results of this pre-test to One Call, to indicate what response could be expected.

In our meeting with One Call we spent most of the time clarifying questions from both sides of the table regarding the implementation of the experiment. We agreed that One Call would handle the division of groups, recording the data and reporting to us; all according to our instructions. Because of the travelling distance from Kristiansand to Oslo, we agreed that we would continue communicating via e-mail and phone, and we were assigned a specific contact person to handle this communication.

The instructions given to One Call during this meeting were as follows:

- Separate the customers aged 18-30 from the rest of the customer base.
- From this group, randomly select five groups for us to use in the experiment.
- Send out each of the nudges by text message to one random group, four in total, while leaving one group "as is". This should be done three days before due date.
- When payments are in for February's bill; register how many paid within due for each of the groups.

- Also record how many customers that pays within due in March for each of the groups.
- Report to us as soon as possible for analysis.

In addition to this, we asked them to extract information from their data systems concerning the total default rates from the selected demographic from January through March of the previous year. This was for the purpose of comparing any possible results with the same time period from 2012.

The representatives from One Call were very enthusiastic about our project as it fitted right into their current plans. The cell phone industry was already working closely with the Norwegian Consumer Ombudsman concerning the issue of young adults defaulting on bill payments. One Call has a clear overrepresentation of younger customers, and wanted to use the findings of our thesis as basis when deciding on the implementation of permanent, long-term measures to increase awareness of payment and payment punctuality among young debtors.

We wanted to test nudging according to relevant theory. A theory is defined by Ghauri and Grønhaug (2010) as *«a set of interrelated concepts, definitions and propositions that present a systematic view of specifying relations among variables with the purpose of explaining and predicting phenomena»* (Ghauri & Grønhaug, 2010, p.37). In the following communication with One Call, we collaborated closely to properly formulate the messages that were to be sent out. After some discussion, we decided to move forward with the following text messages:

Common reminder:

This message did not contain any information beyond a simple reminder, only that there was an upcoming invoice to be paid. This information is also included in the other text messages, to make them as similar as possible. Thereby we can more correctly separate the different effects from the additional information.

«Hei. Vi minner om at denne måneds faktura har forfallsdato XX.02.2013. Du finner kopi av din faktura på Mine Sider på www.onecall.no. Ha en fin dag! Hilsen One Call»

Message based on the anchoring-effect:

By emphasizing the number zero, we hope to anchor the customers' attention in this particular number. It is somewhat similar to the loss aversion-nudge but the effects on decision-making mechanisms are different. By anchoring in the number 0, we aim to make this the default for the customer, and subsequently that the customer will be reluctant to make decisions that would alter the current situation.

«Hei. Vi minner om at denne måneds faktura har forfallsdato XX.02.2013. Betaler du til forfall, får du kr 0 i gebyrer som kan komme ved for sen betaling. Du betaler da kun for ditt abonnement og forbruk. Du finner kopi av din faktura på Mine Sider på www.onecall.no. Ha en fin dag! Hilsen One Call»

Message based on loss aversion:

As people tend to experience a given loss as much worse than the good of gaining the equivalent sum of money, the aim is here to focus on a potential loss. The message is trying to convince the customer that they do not stand to gain anything by paying their bill on time, but they will *avoid* the extra loss represented by a late-payment fee.

«Hei. Vi minner om at denne måneds faktura har forfallsdato XX.02.2013. Ved å betale denne til forfall, unngår du purring på 63kr. Du finner kopi av din faktura på Mine Sider på www.onecall.no. Ha en fin dag! Hilsen One Call»

Message based on social norms:

In accordance with the theory on social nudging, we developed a message that plays on customers' social conscience or adhere to social norms. Most people will seek to be just as "good" as everyone else, and avoid sticking out from the group in a negative manner. They might want to be a part of a group that adheres to the same social norms (e.g. paying what you owe on time), or they might think they should act like the majority as this can serve as guidelines for what is normal or "smart".

«Hei. De fleste av våre kunder betaler sin faktura innen forfall og det setter vi stor pris på. Vi minner om at denne måneds faktura har forfallsdato XX.02.2013. Du finner kopi av din faktura på Mine Sider på www.onecall.no. Ha en fin dag! Hilsen One Call»

These text messages represent the different types of nudges that were sent out to our sample of customers. They have been developed according to theory, meaning that they are in accordance with requirements to count as a mere nudge as explained in the theory. Later they have been customized to fit One Call's "tone of voice" towards their customers. The wording is meant to be as similar as possible to extract the pure effect of different information while eliminating redundant "noise". In addition to the four groups, we also kept one control group to separate the effect of the nudges from the effect of any other unknown factors.

The experiment was carried out as follows: Text messages were sent to the four treatment groups three days before due date on February's bill. We chose to send out the text messages relatively close to due date so that customers would not forget the message, but not so close that the effect could not be observed. A couple of days after the due date, One Call registered how many customers in each group had paid on time or not. They also recorded payment rate from the control group with the same number of subjects as the treatment groups. Those who had not paid would receive a late-payment reminder with an additional fee. One Call also kept track of the sample groups' payments in March, so that we could see if the effect of nudging was also prevalent in the following month. After collecting all the data, One Call sent us the primary data for us to analyze.

We have had some challenges with communication during the process. We have to acknowledge that we might have underestimated the importance of being specific. In the spirit of being polite, we might have caused our fair share of confusion and misguidance. Also, it proved challenging to explain something we know so well to someone who knows so little about it. Adding to this, we had to deal with several contacts within the organization that all had different levels of knowledge about our project.

We did not have direct contact with the system administrator who was responsible for the actual data collection. So the communication process itself consisted of a number of links, all subject to cause confusion. This added to the time span of catching miscommunication in the process and resolving any confusion.

We first received a set of data from all months requested, where the records seemed to be in accordance with what we had initially expected. But the data from March seemed to be off. All

the data collected from March showed a remarkable lower level of defaults compared to any of the other months from which we had received data. One would expect that the messages would have had the most influence on February's default rate, and then a slight default increase in March – maybe even close to the default rate for the entire customer base. The default rate from the control group in March showed only a fraction of the total default rate. This made us nervous that something had gone wrong in the data collection process, as they should be quite similar, and made us question whether all the data received was in fact incorrect. After much correspondence back and forward we managed to locate the mistake that had been made, and finally received an updated and correct data set. After this, we also contacted a third person in the organization to check the data independently. This was done to ensure that the data were correct and collected according to the instructions given.

3.8 Reliability and Validity

Reliability refers to the stability of the measure, being an indicator of internal consistency (Zikmund et al., 2010). A measurement is reliable when several attempts to measure the same phenomenon converges on the same result (Zikmund et al., 2010). You would expect that if you measured an individual according to one variable five times, you would get the same result every time. Reliability is all about consistency, or the consistency of measurement. If the measurements lack reliability, we tend to get inconsistent measurements resulting in false or spurious findings (Kahneman & Tversky, 1979; Zikmund et al., 2010). We therefore need to secure that the measurement process is performed without error.

We distributed our text messages containing nudges only once; just before due date in February of 2013. Then, we collected data concerning actual payment ratios for January, February and March for both 2012 and 2013. By doing this, we were able to record the isolated effect of the treatments, and compare the results to the default rates from one year ago. Our unit of measurement was straight-forward; do the customers pay on time or not. One Call has accurate tracking systems and software, making it easy to register and extract data on the portion of defaulted payments from our sample. As we have not had the opportunity or the resources to go through the data ourselves, we rely on the accuracy of reports from our contact person in One Call. This means that for the time span we are gathering data from, we accurately measure the percentage that did or did not pay on time - every time. Therefore, we stipulate that our

experiment and our thesis have strong reliability and that we will not encounter problems related to inconsistent measurements.

Validity is the extent to which a score truly represents a concept (Zikmund et al., 2010).

Internal validity exists only if the independent variable is in fact responsible for the variance in the dependent variable (Zikmund et al., 2010). In other words, do we really measure what we intended to measure? If the findings were influenced or confounded by other factors, the conclusions made based on the study might not be valid. In this study we have clearly defined what it is we want to measure. As we have collected primary data, there is little room for "noise" to interfere with what we have actually measured. We used a control group to check that we measured the effect of our nudges specifically, separating the effect from other factors and spurious relations. We are therefore confident that the internal validity of our experiment is satisfying.

External validity is the accuracy to which the findings may be generalized beyond the specific study (Zikmund et al., 2010). If the experiment was performed by using a different sample or population, would the results be the same? In our study we have investigated a randomized sample from One Call's customer base aged 18-30. Based on the method of selection, we firmly believe that our findings can be generalized to this larger group of people with solid external validity. The results are primarily valid to count for the Norwegian cellular industry, on the given age demographic. The question remains whether we can generalize the results to count for other age groups, other industries, and outside the national borders of Norway.

4. Data analysis

In this section we present and analyze the data material collected during our experiment.

4.1 Results

On the next page is a table presenting the raw data material. Our primary data was collected from January through March of this year, 2013. From this year we collected both total default rates within the selected age demographic and specific default rates in our treatment groups and the control group, from each month. Ideally, we would also have benefited from having data from April of this year, to further establish any possible longer term effects. As this was not part of the initial agreement with One Call, these data have not been registered and we have therefore not been able to get hold of them. We did, however, acquire total default rates from January through March of 2012 as added information to the findings from this year, adding to the discussion of our findings.

Table 4.1: Original data

Customers aged 18-30	Total	Non-default	Default	Default rate
January of 2012	50650	47023	3627	7,161 %
February of 2012	51124	47530	3594	7,030 %
March of 2012	53196	48604	4592	8,632 %
January of 2013	82292	74322	7970	9,685 %
January 2013 *estimated*: Reminder	500	452	48	9,685 %
January 2013 *estimated*: Anchoring	500	452	48	9,685 %
January 2013 *estimated*: Loss aversion	500	452	48	9,685 %
January 2013 *estimated*: Social norms	500	452	48	9,685 %
January 2013 *estimated*: Control group	500	452	48	9,685 %
	2500	2258	242	9,685 %
February of 2013	86486	78240	8246	9,534 %
February 2013: Reminder	500	482	18	3,600 %
February 2013: Anchoring	500	472	28	5,600 %
February 2013: Loss aversion	500	465	35	7,000 %
February 2013: Social norms	500	436	64	12,800 %
February 2013: Control group	500	445	55	11,000 %
	2500	2300	200	
March of 2013	88060	77469	10591	12,027 %
March 2013: Reminder	500	464	36	7,200 %
March 2013: Anchoring	500	435	65	13,000 %
March 2013: Loss aversion	500	449	51	10,200 %
March 2013: Social norms	500	429	71	14,200 %
March 2013: Control group	500	437	63	12,600 %
	2500	2214	286	

January through March of 2012

In January 2012 the default rate in the total customer base ranged 18-30 year old, was 7,16 %. In February 2012 the default rate was slightly reduced to 7,03 %. In March the default rate was 8,63 %, which is a noticeable increase from the two previous months.

January of 2013

In January of 2013 the default rate for the total customer base between 18-30 years of age was 9,69 %. Since our project had not yet started at this point in time, One Call did not register the specific default rates of our treatment groups at this stage. However, since our sample is assumed to be fully randomized and has not received any treatment in January, we estimated that the default rate in each group could be expected to be the same as for the total demographic, namely 9,69 % for each group of 500 units.

February of 2013

In February our sample groups were subjected to treatments, making this month the first point in time where we could record any possible effects of our experiment. The total default rate in our chosen age demographic was 9,53 %, while the control group of 500 units had a default rate of 11 %. In the sample group that received only a reminder, only 3,60 % defaulted their bill payment. In the group that received the message based on the anchoring-effect, 5,60 % defaulted. The group that received the message based on loss aversion had a default rate of 7,00 %. All of these were lower than the default rates in both the total demographic and our appointed control group. Interestingly, the group that received the message based on social norms had a default rate of 12,80 %, ranging higher than the total default rate and the recorded default rate in the control group in this month.

March of 2013

We also recorded whether the treatments could have an effect on the default rates in the following month, supporting a possible longer term effect, without subjecting the groups to a new series of treatments. In March the default rate for the total demographic was 12,03 %, while the control group had a default rate of 12,60 %. Both of these showed a slight increase from February. One would therefore expect all the treatment groups to have slightly increased default rates. We found that the group that received only the reminder had a default rate of only 7,20 %,

almost half the default rate recorded for the control group. The other groups showed default rates that are more in line with what one would expect; The group that received the message based on the anchoring-effect had a default rate of 13,00 %, the group that received the message based on loss aversion had a default rate of 10,20 %, and the group that received the message based on social norms had a default rate of 14,20 %. All of these rate noticeably higher than in the previous month.

2012 compared to 2013

During the last year, One Call has recorded a general increase in total default rates. There was also a substancial increase in the total customer base aged 18-30. Based on the data material, we were also able to spot possible seasonable fluctuations; in both years March had a noticeble higher default rate than in the previous two months of the year.

4.2 Analysis

The aim of our experiment is to investigate whether nudging may reduce the number of young cell phone customers that default on their bill payments. As we have reported the data collected in the above section, we will now move on to the analysis. While there are several ways to analyze quantitative data, we have chosen to use descriptive methods supported by statistical methods for testing our hypotheses. This is sufficient to establish any relationships between the nudges and the default rates, as our original data is fairly straight-forward in nature.

The analysis is based on calculations performed using SPSS (*Statistical Package for the Social Sciences*). The results from our original calculations can be found in the appendix.

4.2.1 Descriptive statistics

Descriptive statistics are used to describe the basic characteristics of a data collection. We cannot draw any conclusions beyond what can actually be observed from the data material, as the aim is only to describe and structure the main content of our data.

In this paragraph we summarize the findings within our sample groups. Measures of central tendency and measures of variability are summarized in table 4.2 and table 4.4, representing measures from February and March of 2013, respectively. The differences between the groups can easily be observed when the data is displayed graphically, as shown in figure 4.1 and figure

4.2, corresponding to the aforementioned tables. We have further set up tables telling the percentage difference between each treatment group compared to the control group of the corresponding month.

February

Table 4.2: Summary of measures, February 2013

Treatment group:	Reminder	Anchoring	Loss aversion	Social norms	Control group
Sample size	500	500	500	500	500
Mean	0,036	0,056	0,070	0,128	0,110
Variance	0,035	0,053	0,065	0,112	0,098
Standard deviation	0,186	0,230	0,255	0,334	0,313

These are the measures used when performing the independent samples t-test for February to verify or reject our null hypotheses in the following paragraph. We observe the differences in means and variances between the groups.

Figure 4.1 displays the differences in default rates between subject groups in February. The separate bars indicating default rates for each group show a noticeable reduction in default rates in three of the four treatment groups; the group that received the common reminder, the group that received the message based on the anchoring-effect, and the group that received the message based on loss aversion. In the last treatment group, that received a message based on social norms, the nudge seems to have had the opposite effect of what we expected. The default rate in this particular group is higher than in any of the other groups, including the control group. This indicates that not only did we not get the desired effect, but the effect recorded was worse than what could be expected if the group was not subjected to any treatment.

Figure 4.1: Default rates in February, 2013
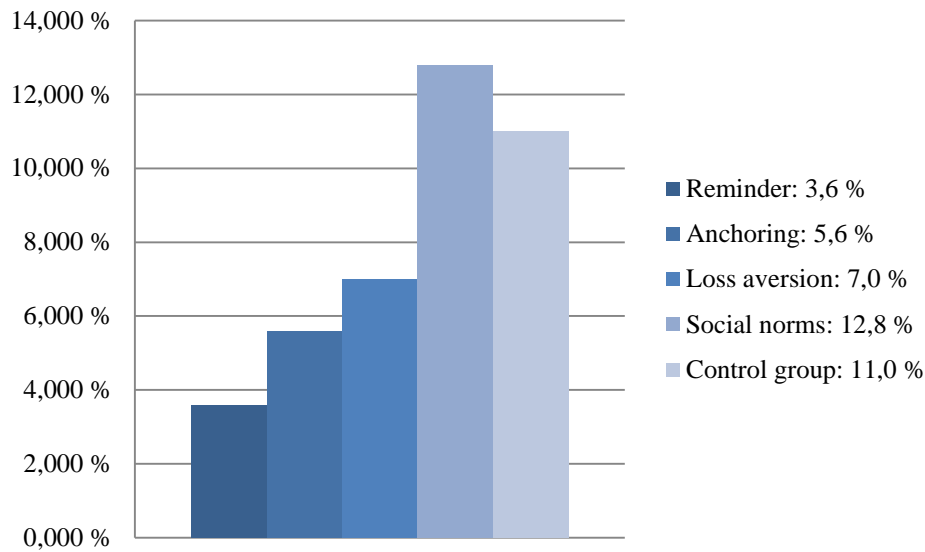
In table 4.3 we have calculated the percentage difference between each treatment group and the control group. The default rate in the control group was approximately three times the default rate found in the group that received a reminder, about twice as large as the group that received the anchoring-nudge, and 57 % larger than the default rate in the group that received the loss aversion-nudge. These numbers suggest that the three aforementioned nudges did in fact influence people's actual behavior, and reduced the number of defaults. The effect of nudging based on social norms is an interesting finding. The default rate is actually 16,36 % *higher* in this group compared to the control group.

Table 4.3: Percentage difference in each treatment group measured against the control group, February 2013

Treatment group:	Reminder	Anchoring	Loss aversion	Social norms
% difference compared to the control group	(-205 %)	(-96,43 %)	(-57,14 %)	16,36 %

March

We also wanted to check whether the effects of nudging in February could transfer beyond the initial treatment period. To test this, we also recorded data from the sample groups in March. By doing this we are able to compare the results between the two time periods, and determine whether the effect is indicated to be prominent on a longer term basis.

Table 4.4: Summary of measures, March 2013

Treatment group:	Reminder	Anchoring	Loss aversion	Social norms	Control group
Sample size	500	500	500	500	500
Mean	0,072	0,122	0,094	0,142	0,126
Variance	0,067	0,107	0,085	0,122	0,110
Standard deviation	0,259	0,328	0,292	0,349	0,332

These are the measures used when performing the independent samples t-test for March to verify or reject our null hypotheses in the following paragraph. We observe the differences in means and variances between the groups.

Figure 4.2 displays the differences in default rates between the subject groups in March. From this graph it seems that it is the reminder and the nudge based on loss aversion that could have a longer lasting effect, with noticeable lower default rates than the control group. The other sample groups seem to be somewhat returning to a normal state with default rates that are similar to the control group.

Figure 4.2: Default rates in March, 2013
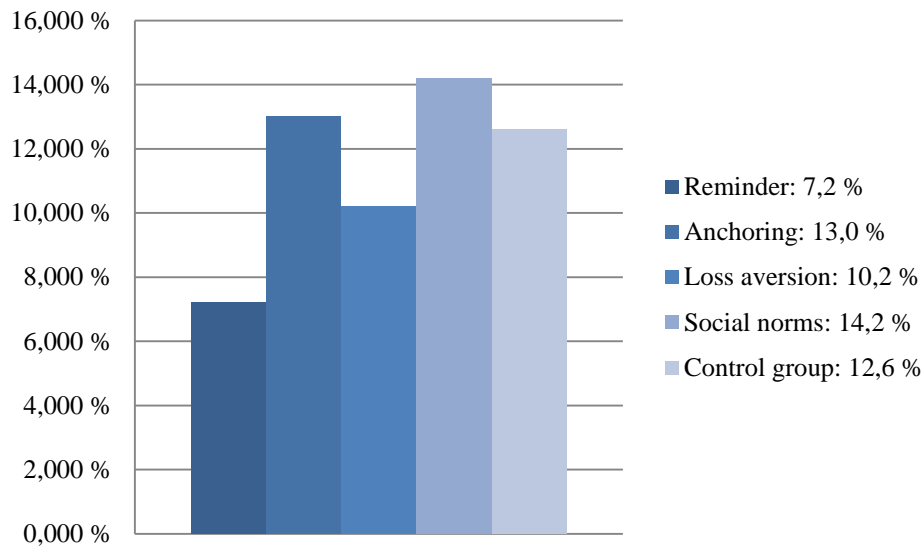
From the calculations presented in table 4.5, it seems that neither the message based on the anchoring-effect nor the message based on loss aversion has a significant effect in the month following the initial treatment. These smaller differences from the control group might be explained simply by natural variations. It is interesting, however, that the group that received the message based on social norms, still has a default rate that is higher than the control group. It might not be of significance, but it is still interesting that this deviation from our expectations would be prominent in the month following the treatment period.

Table 4.5: Percentage difference in each treatment group measured against the control group, March 2013

Treatment group:	Reminder	Anchoring	Loss aversion	Social norms
% difference compared to the control group	(-75 %)	3,1 %	(-23, 5 %)	12,7 %

4.2.2 Hypothesis testing

Our hypotheses reflect our problem definition; we wish to examine if either of the messages will reduce default rates to be significantly lower in the nudged groups compared to the control group. To determine if our findings support our assumptions, we test the hypotheses to conclude whether we can reject the null hypotheses we have put forward.

The most commonly used procedure for testing hypotheses requires that we directly test our null hypothesis, to conclude that we can either reject or fail to reject our null hypothesis (Triola, 2011). Based on the data material we have gathered we will conduct statistical tests to uncover if the groups that received nudges have significantly lower default rates than the control group of the corresponding month. We use these statistical methods to support our descriptive analysis as presented above, and draw final conclusions. By using statistical methods to test our hypotheses we may uncover information and results that may not be intuitive by simply describing the data collected.

When we conduct a hypothesis tests we are at risk of making two types of errors. We may reject a null hypothesis that is actually correct, which is called an error of the first kind. The other mistake is failing to reject a null hypothesis that is false, which is called an error of the second kind. The probability of rejecting a true null hypothesis is denoted as the significance level. We test our hypotheses on a significance level $\alpha = 0{,}05$, which means that we can expect to reject the null hypothesis in 5 % of the instances where it is actually correct. Another way of looking at this is to say that we will correctly reject the null hypothesis with 95 % certainty (Triola, 2011).

The T-test

A t-test is a statistical test that can be used to determine if the means of two sets of data are significantly different from each other (Triola, 2011). We want to test whether or not the results from each of the nudged groups are significantly different from the control group, for each month of collected data. We use the independent samples t-test, as our sample groups are separate sets of independent and identically distributed samples, with different variances (Triola, 2011). When looking for a possible prolonged effect, we use a paired samples t-test because we are measuring the same treatment group in two different time periods. This test will be used to prove whether or not the results found in the month of treatment are directly transferred to the following month. If

we find is a significant difference between the results from a treatment group in February and the results from the same group in March, we cannot claim that the effect of the treatment is directly transferable to the next month.

Due to the nature of our research question and hypotheses, we test our hypotheses in one direction when performing the independent samples t-test. We wish to examine whether nudging may have reduced defaulted payments, and test if the default rates in the nudged groups are significantly *lower* than in the corresponding control group. Therefore we have chosen to do a one-tailed t-test, so we are testing for a relationship in one direction (Triola, 2011). The independent t-tests will be structured as followed:

Null hypothesis: H_0: $\mu_1 = \mu_2$

Alternative hypothesis: H_A: $\mu_1 > \mu_2$

Test statistic: $\dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_1{}^2}{n_1} + \dfrac{s_2{}^2}{n_2}}}$

Data: We have $n_1 = n_2$ for observations of x_1 and x_2. In our case x_1 is the observed default rate in the group that did not received treatment, and x_2 is the observation of the default rate in the group that did received treatment.

The results from performing the t-test is measured against significance level α to see if there are significant differences between the two groups, and by measuring the test statistic against a critical t-value to determine whether or not we are able to reject H_0.

Conclusions that can be drawn on the basis of this test: Reject H_0 if: $|t| > t_\alpha$ *or* $p < 2\alpha$

Reject H_0 if: $t < -t_\alpha$ *or* $p < \alpha$

Reject H_0 if: $t > +t_\alpha$ *or* $p < \alpha$

T-test localization

H_0: $\mu_c = \mu_i$ No difference in default rates in treatment group compared to control group

H_A: $\mu_c > \mu_i$ The default rate is lower in treatment group measured against the control group

i = {1,4}

for m = F, M

This is the general hypothesis that will be used to test four different treatment groups, in two different time periods, meaning that we will perform a t-test for each treatment group in each month from which we have collected data. The H_0 and H_A will be the same for each separate treatment group.

We can make inferences about our hypotheses by looking to p-values and test statistics, measured against the significance level and the critical t-value. The critical t-value that we will measure our test statistic against is found by using a t-distribution table. The critical t-value is calculated on the basis of the selected significance level, and the degrees of freedom in our samples. SPSS only has the option of conducting a two-tailed independent samples t-test. Results from a two-tailed test are measured against different critical t-values than one-tailed tests. The critical t-value we have used has therefore been found using a t-distribution table. We have selected to test our hypotheses in one direction, at a significance level $\alpha = 0, 05$ and our samples have 500 (or 500 - 1 = 499) degrees of freedom. This provides us with a critical t-value (t_α) of 1,648.

Testing our findings from February

Table 4.6 summarizes the t-values and p-values found by conducting a t-test between each treatment group and the control group in February.

Table 4.6: Results from t-test between treatment groups and control group, February 2013

	T-value	**p-value**
T-test Reminder vs Control	$-4{,}539 < -t_\alpha$	$P < 0{,}001 < \alpha_{0,05}$
T-test Anchoring vs Control	$-3{,}017 < -t_\alpha$	$P = 0{,}002 < \alpha_{0,05}$
T-test Loss aversion vs Control	$-2{,}213 < -t_\alpha$	$P = 0{,}027 < \alpha_{0,05}$
T-test Social norms vs Control	$0{,}878 \ > -t_\alpha$	$P = 0{,}38 > \alpha_{0,05}$

From the information conveyed in this table, we are able to make inferences on whether or not we can reject our null hypotheses for February.

We see that for each of the treatment groups noted as reminder, anchoring and loss aversion, the p-values calculated are lower than the significance level that we selected for our test. This indicates that the default rates for these groups are significantly different from the control group at a significance level of $\alpha = 0,05$. Adding to this, the aforementioned treatment groups have t-values that are lower than the critical t ($t < -t_\alpha$). This empirical evidence leads us to conclude that we can reject H_0 for these treatment groups. We have found grounds to state that the default rates for these treatments groups are significantly lower than the default rate in the control group. Our nudges seem to have served their purpose. As could be expected, the t-test reveals that the treatment group that received the message based on social norms, is neither significantly different from the control at $\alpha_{0,05}$ nor is there grounds for rejecting H_0 as the t-value is not smaller than the critical t-value ($0,878 >- t_\alpha$). We have not found evidence that this nudge has lowered the default rate.

Testing our findings from March

Table 4.7 summarizes the t-values and p-values found by conducting a t-test between each treatment group and the control group in March.

Table 4.7: Results from t-test between treatment groups and control group, March 2013

	T-value	p-value
T-test Reminder vs Control	$-2,868 < -t_\alpha$	$P = 0,004 < \alpha_{0,05}$
T-test Anchoring vs Control	$0,189 > -t_\alpha$	$P = 0,850 > \alpha_{0,05}$
T-test Loss aversion vs Control	$-1,194 > -t_\alpha$	$P = 0,223 > \alpha_{0,05}$
T-test Social norms vs Control	$0,742 > -t_\alpha$	$P = 0,458 > \alpha_{0,05}$

We can see that the results from testing our data are very different in March when compared to February. The data conveyed in table 4.7 tells us that only the treatment group that received the reminder had a significantly different default rate from the control group in this month. It was also the only group that produced a t-value that was lower than the critical t-value. This means that we are able to reject H_0 only in the case of the group that received the common reminder, suggesting that this particular nudge had a longer term effect. In the section on descriptive statistics we noticed that loss aversion might also have a longer term effect due to observable

lower default rates than the control group. By performing the t-test we see that this relationship was not significant, and we are unable to reject H_0 for this treatment group. We have found evidence supporting that is was only the common reminder that produced default rates indicating a prolonged effect in the month following actual treatment.

Comparing our findings for each treatment group between February and March

Even though it seemd fairly clear that the messages in general did not have a significant lasting effect on the treatment groups, we performed a paired samples t-test to examine to what degree the results differed in each group between February and March. If there is a significant difference between the results obtained in a tretment group in February and the results obtained from the same treatment group in March, we can at least provide evidence that the effect of the treatment is not the same, or close to similar, in both months.

Table 4.8: Results from t-test between February and March for each treatment group

	p-value
T-test Reminder February vs Reminder March	$P < 0{,}001 < \alpha_{0,05}$
T-test Anchoring February vs Anchoring March	$P < 0{,}001 < \alpha_{0,05}$
T-test Loss aversion February vs Loss aversion March	$P < 0{,}001 < \alpha_{0,05}$
T-test Social norms February vs Social norms March	$P < 0{,}001 < \alpha_{0,05}$

By conducting a paired samples t-test we found that with 95 % certainty there was a statistically significant difference between the results from February and March, for each of the treatment groups. This means that any effects found on default rates in February did not directly transfer to the default rates in March.

5. Conclusions

In this section, after having presented, reviewed and statistically tested our data material, we are able to make some inferences about the study performed.

5.1 Discussion of findings

The results from February indicated that three of our messages had the effect of significantly lowering default rates on the bill following the treatment. This is in line with previous research and what we had initially anticipated. The treatment groups that received the common reminder, the message based on the anchoring-effect and the message based on loss aversion, all had default rates that were significantly lower than the control group in February. We stipulated that it would be possible to influence people's decision-making process by introducing such external stimuli. As projected, our findings indicate that nudging using text messages as the external stimuli has served to influence people's decision-making process in the form of more people choosing to pay their bills on time. With that being said, the actual effect of the messages on default rates differed. This difference in observed effect may be explained by the differences in content of the text messages.

The common reminder had the best results in terms of lowering the default rate, indicating a 200 per cent difference compared to the control group in February. There might be several reasons for why this particular nudge had such a positive effect. The most intuitive reason for any positive effect of this nudge would be that simply reminding customers of an upcoming bill, shifts people´s attention from other tasks or spendings to ensure that the particular bill is paid, thereby generating more people to pay within due date. Here, we might have captured those who simply forget that they have received a bill to be paid and the upcoming due date, or those who may otherwise choose to spend their last money of the month on other things. But, as the other messages also contained a reminder, this explanation is insufficient in explaining why this particular message had a better effect on the default rate than the other messages that also proved effective in reducing defaulted payments. When compared to these two other messages, the pure reminder was very short and concise, which increases the probability of the message being read and fully understood. Also, by leaving out any additional information, i.e. any other type of influence, the message may be perceived as a mild "slap on the wrist", giving the impression that

the company perceives you, or their customers in general, to be likely not to pay on time. This may in turn make people want to "improve" their behavior and/or prove the company wrong. This reasoning might also help explain why this specific treatment group had the lowest default rates in March as well. The message might have felt unpleasant to receive, and is therefore likely to be remembered and acted on in the future. By continuing to pay the bill on time, the customer ensures that the message will not be repeated in the future.

All the text messages that were distributed contained a reminder that the customer had an outstanding bill that was due within a few days. We had anticipated that the nudges based on heuristics and biases, i.e. the text messages with additional information to the pure reminder, would have a stacking effect; enhancing the effect of the pure reminder and maybe even give a synergy effect. This proved not to be the case as these messages produced slightly higher default rates than the simple reminder, though lower than the control group. All the reminders should, in theory, equally serve to remind the customers about the upcoming due date. We expected that the additional information would influence the customer to a greater extent than the reminder alone, producing an even lower default rate than the reminder-nudge. The message based on the anchoring-effect and the message based on loss aversion did produce significant results, indicating that they had the expected effect, but still produced higher default rates than the group of customers that received the common reminder. This discrepancy in results may stem from different reasons. One possible explanation is that these messages influenced the decision-making process in different ways than the simple reminder. The message focusing on an added fee if defaulting payment, highlighted the cost side of the issue. This could be an effective way to influence those that were not already familiar with the actual consequences of defaulting on bills. The message that focused on the anchor, informed people that paying within due is the only way to avoid paying for services they have not made use of. This might have captured those individuals that are conscious about not wasting their money. So, compared to the reminder, we might have influenced different aspects of the payment decision. Regardless, it was interesting to see that these effects did not add to the reminder-effect, and resulted in *fewer* people paying on time than in the group that received the reminder. We expected that the messages developed according to decision-making theory would capture individuals based on heuristics and biases *as well as* influencing those individuals that were prone to forget their financial obligations. One explanation for the discrepancy in results may be that the presence of additional information may

have changed the point of focus in the messages. It is a possibility that the additional information reduced the effectiveness of the reminder by leading the receiver to be overly, or only, focused on the additional text. We may have been able to influence the individuals as described above, but failed to influence those who simply needed to be reminded. It may also be the case that the messages were simply too long. Most modern cell phones show a preview of a received text message, most of the times in full, making it impossible to ignore. We are relatively certain that most people will read the text messages they receive, but if entire message is not shown in the preview, there is a possibility that test subjects have refrained from opening and reading the entire message, overlooking the content of the continuous text.

The results found in the group that received the message based on social norms were unexpected, and thereby in itself interesting. We had expected to observe a clear reduction in default rates compared to the control group. This expectation was based on previous studies and theory, in addition to the pre-test conducted before starting this project. Combined, this knowledge made us anticipate that this was the group in which we would find the best results. As our experiment strongly resembles Allcott's study from 2011; successfully reducing people's energy consumption using nudging based on social norms, as well as the study performed in 2012 on tax compliance in Britain, we had expected to find that this nudge would noticeably lower default rates for bill payments. Instead, we found the opposite effect to be present in both months of recording payments. The appointed group that received this nudge had a *higher* default rate than the group that was not subjected to any treatment, both in February and in March. Based on our findings we do not have grounds to claim that nudging based on social norms does in fact *increase* default rates, but we cannot eliminate the possibility of this negative effect being a direct result of our nudge in this specific case. We have tried to find explanations for why this nudge proved ineffective. One possibility is that the text message was too long, as it was the longest message sent to the treatments groups. As mentioned above, on most modern cell phones the message will appear on the screen without it physically being opened. It is possible that this message was too long to initially fit the screen. This could mean that people did not bother to open and read it properly, or simply disregard the content due to its inconvenient length. This should lead to a default rate quite similar to the control group. The fact that the results showed a higher default rate for this group than in the control group, suggests that the message might have discouraged some of the customers from paying on time. It is possible that the message was

perceived to be intrusive. But, as people are not likely to "get back" at the company by not paying their bill on time, a more probable explanation for our findings is that people might have misinterpreted the content of the message. We had hoped that people would respond by wanting to act in consistency with an accepted social norm, and act according to what the majority of people perceive to be the correct decision. We might, however, have ended up with sending a signal that most people pay on time so that it's quite all right for the specific customer not to. If this was the case, our nudge was incorrectly formulated, did not serve its intended purpose, and thereby indicated to be counter-effective. It might also be the case that nudging based on social norms has a different effect in writing than it would have if communicated in person; nudging based on social norms might require a physical social setting. The explanation for the discrepancies between our expectations and our findings may also be a result of differences in contexts and demographics; the experiments we have previously presented were conducted in a different context and studied the effect of social nudging on a different demographic than ours. The demographic in these experiments was mostly settled adults at a later stage in the life cycle, whereas our demographic is young adults. This is supported by the findings of Gulbrandsen (1999); if propensity to default is a result of people´s moral, and morality is continuously developed throughout the life cycle, this could help to explain the relatively high default rates in this particular group. It might also be that young adults have a lesser need to be part of the socially accepted group in terms of bill payments.

When looking for a longer term effect of the nudges, we found that most of our successful text messages lost their effect in March. We were only able to find indications of a prolonged effect on default rates in the group that received the common reminder, though this effect was also decreasing compared to the previous month. The results from the statistical tests showed that there was a significant difference between the results in February and the results from March, for all treatment groups that showed lower default rates in February. This implied that the results found in February did not directly transfer to March for any of the nudges. As mentioned, the group that received the pure reminder had a significantly lower default rate compared to the control group also in March. But the effect did not directly transfer by being equal in both months. This fact indicates that people would need repeated nudges to eliminate a sporadic effect. This may simply be due to people's forgetfulness and the short-term impact these text messages give. The reminder could easily be repeated, but the other nudges would have to be somewhat

varied not to seem forced. On the other hand, if these text messages were to be distributed on a regular basis, one would expect a decreasing effect due to the customers' familiarity of the content. Repeating the nudges could be perceived as nagging, not nudging.

We also wish to comment on which individuals we believe to have successfully nudged. We have previously presented Brusdal and Berg's characterization of how young adults with financial difficulties can be divided into four categories based on behavioral traits. The behavioral traits of individuals in each of our treatment groups may have affected the results of our study. We believe to have successfully nudged individuals falling under the "absent-minded" or the "shopper" categories. The individuals portrayed as being "absent-minded" are generally forgetful and do not have adequate overview of their financial situation. These individuals may have been affected by the reminder as they clearly could benefit from receiving one, and the message based on the anchoring-effect as this message contains information about the added fee that they may not have known or forgot. The "shopper" is typically shortsighted and lacks impulse control. The message based on loss aversion or the message based on the anchoring-effect may have produced a positive response from these individuals. By making them aware of the cost of defaulting payment, or that they will pay more than necessary by not paying on time, we believe that this information is taken into consideration when prioritizing where to spend their money. It is doubtful that we have been able to influence individuals belonging to the "victim" or the "juggler" categories. People with these traits typically do not take responsibility for their own actions, by blaming someone else or purposely cheating the system. It is unlikely that our nudges would affect the decision-making process of such individuals as they do not accept their own role of the problem. With this being said, we do not have any evidence to support which type of individuals we have successfully affected, and cannot further comment on the distribution of behavioral traits within our treatment groups.

Other remarks to our findings

We also have comment on the fluctuations in default rates in the control group and in the total demographic. Had our findings been ambiguous when comparing the treatment groups to the control group of each month, we would have had to further investigate these fluctuations. It is normal to have a slight difference in default rates from one month to the next, and from one group to another. As the differences were so small that they could be explained by natural

fluctuations, we did not find a need to investigate this any further. We also recorded a slight increase in default rates in March, for both 2012 and 2013. This has been explained by One Call to be normal, as there are several public fees and payments that are due during this particular month of the year. Cell phone bills will typically not be given priority until such bills are paid, as cell phone bills are often perceived to be less important. We also registered a general increase in total default rates from 2012 to 2013. This could be explained by two major factors. First, during the last decade there has been a steady increase in financial difficulties, credit-debt and defaults in Norway, especially accounted for by young adults. The increasing default rates for One Call's customer base may be a consequence of a general societal development. Second, we also have to take into account the rapid growth of One Call's customer base aged 18-30, which has almost doubled in one year. The youngest customers are usually the ones with the most severe financial difficulties. A big part of the new customers might be in the group of younger adults, as this is One Call's target market. This might also help explain the general growth in default rates.

We considered using the default rates from the entire population, being One Call's customer base aged 18-30, as the control group. This would have given more accurate picture of default rates for individuals that did not receive treatment. However, we chose to use a control group consisting of 500 people for two reasons: Firstly, it would have been a complicated process to separate the 2000 individuals that received treatment from the entire population. So there was a cost issue in terms of using the population as a whole. Secondly, using the same number of people in all groups simplified the statistical testing of the data material. As the differences in default rates between the total demographic and the control group are fairly small, we considered the differences to be insignificant for the study and the conclusions drawn.

5.2 Limitations

We acknowledge there are several limitations to our experiment. In this section we address some of these limitations and discuss how these affected the study.

When conducting the experiment for our Master's Thesis, we faced two critical constraints; time and resources. It would not be possible to conduct the experiment if we did not get started in time to execute the treatments and collect data material from the following two months. With five months to work on our thesis, time was of the essence. Had we been able to collect data for a longer period of time, both prior to and following the selected time period, we would have been

able to further establish any relations found and tested the data more thoroughly. Also, we have not been able to exclude particular circumstances that might affect our study. Current economical, political and social conditions, in addition to others, may have affected the results. To hedge against such influence, one would have to extensively expan the timeframe of the study. Another constraint was resources - both the collaborating company's and ours. We had to balance the extent of our study against the restraints of being students, and against the need to find a willing company and their possibilities for execution. We also had to adapt and adhere to some of the company's wishes.

We were not able to record payments from individual customers in the experiment, nor do we have any characteristics of the subjects. We have only been able to collect data on an aggregated group level, disregarding individual response to the treatment and whether the effect might have relations to specific characteristics. We chose to not request this information from One Call due to limited time and resources. Had we been able to track the development of each subject, we could have detected the individual response to the treatment over time, and possible relations between specific characteristics and the effect of nudging. Also, we cannot be sure that the subjects who defaulted February's bill are the same individuals who defaulted their payment in March. Had we been able to track individuals, and had more information about them, we would have had a more extensive study and subsequently more detailed results. Due to the lack of individual tracking it has not been possible for us to differentiate between reasons for why people do not pay their bills on time. We are not able to pinpoint whether people actively chose not to pay, forgot to pay, or if they did not have the available funds to pay on time. As we only recorded payments on an aggregated group level, not knowing who pays when or anything about who they are, we cannot make inferences about any specific tendencies for why people typically default on bill payments.

We also need to recognize that the nature of our data material limited the possibilities for statistical analysis. Initially we opted to use regression analysis to analyze the data material, calculating statistical relations and possibly provide grounds for prediction. But the use of such models would not provide any useful information, nor would it help explain our findings. This is partly due to our data being recorded at a group level and partly due to the very limited time span for the experiment.

In terms of the results, our findings might in part be result of a «first-mover effect». The treatment, being text messages, is not yet frequently used to influence people's decision-making process. This means that the effect might have been enhanced simply by being new and unfamiliar to the receiver. If every company sent out such messages for all kinds of tasks and events, one could assume that the effect would be reduced or maybe even loose its effect completely. Another important point in this regard is that the effect might have transfferred onto other bills for which the individual did not receive a text message. Receiving a text message reminding the customer to pay the particular bill, may trigger the memory onto other bills that are perceived to be more important, or due prior to the cell phone bill, possibly leaving the cell phone bill unpaid. We need to recognize the possibility that this could have happened.

5.3 Implications

The results from our study provide new evidence to support theory on the effect of nudging people's decision-making process. Our thesis gives a valid contribution to this growing field of research and hopefully active use of nudging as physiological tool to enhance the quality of people's decisions.

5.3.1 Practical implications

The results demonstrate that nudging may reduce default rates accounted for by young adults, and may in turn help to reduce the financial difficulties individuals in this age group may incur due to defaulted payments. This depends on companies' active use of nudging when collecting payments from people in this age group.

The study performed has used a new approach to overcome problems of payment defaults. The results provide evidence supporting the usefulness of nudging as an addition to current means for payment collection. As we have not studied the effect of our nudges when removing existing collection tools of the industry, we cannot conclude that nudging would work in a setting where added fees and other sanctions are absent. Therefore, our contribution to the practical use of nudging is limited to being only an addition to working regulations.

Distinguishing between people who pay their bills promptly and people who have previously defaulted on their bill payments, may provide for more effective use of nudges in this setting. As we have sent out the text messages to randomly selected samples, we have not been able to

observe the effect of specifically targeting those who struggle to pay on time. By only nudging those who have previously defaulted payments, companies may use nudging theory more effectively in their payment collection process as a tool to reduce default rates.

The results of our study are highly relevant for companies that collect payments. As the common reminder produced the best result, the usefulness of our thesis is not only limited to the companies that collect payment, but also companies that need to remind their customers of upcoming appointments or events. This suggests that nudging through text messages can be used in a broader business setting.

5.3.2 Implications for research

The results from this study have implications for future research on people's decision-making process, personal finances, and the concept of nudging specifically.

In our experiment we have not been able to record individual responses to the text messages. Replicating the experiment with individual tracking would provide data for more in depth analysis of individual responses, producing possible grounds for making predictions of the effect. This would make it possible to define the characteristics of those who respond positively to being nudged, and pinpoint these people in future efforts to reduce default rates. It would also be of great value to other industries to investigate whether our results could be transferrable to a different context. This could provide basis for active use of nudging to collect payments.

Our findings concerning using social norms to influence the decision-making process of individuals is not in line with theory and past research on this phenomenon. The results give reason to doubt that social norms can be used to influence young adults in terms of payment behavior. Before we can conclude on the usefulness of this effect, more studies need to be performed. One way to test whether this type of nudge could be effective on bill payments would be to replicate the experiment using an older age group. It might be that more settled adults, who have different characteristics from our sample, would respond positively to this type of nudge as opposed to the age group used in our experiment. We also recognize that the formulation of the nudge may have failed to serve its purpose. It might be that re-formulating the message would produce a positive effect on lowering default rates.

If we assume that the positive effect of the common reminder on both months can be directly attributed to the nudge, it would be interesting to further investigate the duration of the effect. If it is possible to track the duration of the effect, this may indicate how frequent the company should repeat the nudge to keep default rates to a minimum.

Future research on nudging could study the effect of repeated treatment; either the effects of repeating the same type of nudge several times, or testing if there can be effects found from subjecting individuals to different types of nudges over time. We have not been able to uncover past research on the interaction effect of different nudges, so this would be an especially interesting phenomenon to study further.

5.4 Conclusion

The research question for this thesis asked if nudging young adults may reduce their propencity to default on cell phone bills. The empirical evidence put forth through our analysis suggests that nudging using simple and inexpensive text messages has the potential to significantly lower the default rates accounted for by this age group. Our findings thereby support Thaler and Sunstein's theory that nudging the decision-making process may help people make better choices for themselves. If nudging was to be actively used by companies dealing with payment collection, this method of positive influence has great potential to help combat the increasing problems realted to financial difficulties faced by young adults.

Through our study we have uncovered the importance of the information convayed in terms of written nudges. The content of the messages distributed in our experiment produced different results depending on the information included in the message, in terms of lowering default rates for bill payments. We have not been able to find empirical evidence supporting a positive effect from basing our message on accepted social norms. We did, however, find significant results from using a common reminder, a message based on the achoring-effect, and a message based on loss aversion, on the bill following treatment.

In the context of our study, we have found evidence supporting that nudging people's decision-making process will produce positive, but relatively short term effects on young adults´ payment behavior. To sustain the effect, people seem to need repeated nudging over time not to fall back into old patterns, and maybe form improved and longer lasting habits. Allthough we are not able to suggest at what frequency companies should make use of nudging, our findings did reveal a reduced effect over time from our nudges on lowering default rates.

References

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics, 95*(9-10), 1082-1095. doi: 10.1016/j.jpubeco.2011.03.003

Andersen, Otto. (2012). *ME-407: Lecture notes*. Universitetet i Agder. Kristiansand.

Angrist, Joshua, & Lavy, Victor. (2009). The Effects of High Stakes High School Achievement Awards: Evidence from a Randomized Trial. *The American Economic Review, 99*(4), 1384-1414. doi: 10.2307/25592513

Baron, Jonathan. (2008). *Thinking and deciding* (Fourth ed.). Cambridge: Cambridge University Press.

Brusdal, Ragnhild, & Berg, Lisbet. (2011). *Unge vokse og lånefinansiert forbruk* Vol. Nr 4-2011. Retrieved from http://www.sifo.no/files/file77849_oppdragsrapport_4-2011_web.pdf

CAB 099-12. (2012). *Applying behavioural insights to reduce fraud, error and debt*. London: Behavioral insights team: Cabinet Office, UK Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/60539/BIT_FraudErrorDebt_accessible.pdf.

Calzolari, Giacomo, & Nardotto, Mattia. (2011). *Nudging with information: A randomized field experiment on reminders and feedback*. Available at SSRN 1924901.

Charness, Gary, & Gneezy, Uri. (2009). Incentives to Exercise. *Econometrica, 77*(3), 909-931. doi: 10.3982/ECTA7416

Disdier, Anne-Célia, & Marette, Stéphan. (2013). Globalisation issues and consumers' purchase decisions for food products: evidence from a laboratory experiment. *European Review of Agricultural Economics, 40*(1), 23-44.

Edwards, W., Lindman, H., & Savage, L. J. (1963). Bayesian statistical inference for psychological research. *Psychological Review, 70*(3), 193-242. doi: 10.1037/h0044139

Ghauri, Pervez, & Grønhaug, Kjell. (2010). *Research Methods in Business Studies* (Fourth ed.). Essex: Pearson Education Limited.

Grove, William M. (2005). Clinical versus statistical prediction: The contribution of Paul E. Meehl. *Journal of Clinical Psychology, 61*(10), 1233-1243. doi: 10.1002/jclp.20179

Gulbrandsen, Lars. (1999). NOVA rapport 3/99: Å betale for sent. Dårlig råd eller dårlig moral? Retrieved from http://www.nova.no/id/2093.0

Hustvedt, Gwendolyn, & Bernard, John C. (2010). Effects of social responsibility labelling and brand on willingness to pay for apparel. *International Journal of Consumer Studies, 34*(6), 619-626. doi: 10.1111/j.1470-6431.2010.00870.x

Jacobsen, Dag Ingvar. (2005). *Hvordan gjennomføre undersøkelser? Innføring i samfunnsvitenskapelig metode* (Second ed.). Kristiansand: Høyskoleforlaget.

Johnson, E. J., Weber, E. U., Shu, S. B., Fox, C., Dellaert, B. G. C., Goldstein, D. G., . . . Wansink, B. (2012). Beyond nudges: Tools of a choice architecture. *Marketing Letters, 23*(2), 487-504. doi: 10.1007/s11002-012-9186-1

Kahneman, Daniel, Gilovich, Thomas, & Griffin, Dale. (2002). *Heuristics and biases : the psychology of intuitive judgment* (First ed.). Cambridge: Cambridge University Press.

Kahneman, Daniel, & Tversky, Amos. (1979). Prospect Theory : An Analysis of Decision under Risk. *Econometrica, 47*(2), 263-292.

Kahneman, Daniel, Tversky, Amos, & Slovic, Paul. (1982). *Judgment under uncertainty : heuristics and biases*. Cambridge: Cambridge University Press.

Karlan, Dean, McConnell, Margaret, Mullainathan, Sendhil, & Zinman, Jonathan. (2010). *Getting to the top of mind: How reminders increase saving* (No. w16205). National Bureau of Economic Research.

Kirsten, Howard, & Glenn, Salkeld. (2009). Does Attribute Framing in Discrete Choice Experiments Influence Willingness to Pay? Results from a Discrete Choice Experiment in Screening for Colorectal Cancer. *Value in Health, 12*, 354-363. doi: 10.1111/j.1524-4733.2008.00417.x

Menegaki, Angeliki N., Mellon, Robert C., Vrentzou, Anna, Koumakis, George, & Tsagarakis, Konstantinos P. (2009). What's in a name: Framing treated wastewater as recycled water increases willingness to use and willingness to pay. *Journal of Economic Psychology, 30*(3), 285-292. doi: 10.1016/j.joep.2008.08.007

Simon, Herbert A. (1955). A Behavioral Model of Rational Choice. *Quarterly Journal of Economics, 69*(1), 99-118.

Stenseng, Sverre (Producer). (2012). Norsk ungdom har misligholdt 1,1 mrd. *Aftenposten*. Retrieved from http://www.aftenposten.no/okonomi/Norsk-ungdom-har-misligholdt-1_1-mrd-6779085.html#.UUBFKxwieDs

Thaler, Richard H., & Sunstein, Cass R. (2008). *Nudge : improving decisions about health, wealth, and happiness* (First ed.). New Haven: Yale University Press.

Triola, Mario. (2011). *Essentials of statistics*. Boston: Pearson Education Inc.

Zikmund, William G., Babin, Barry J., Carr, Jon C., & Griffin, Mitch. (2010). *Business Research Methods* (Eighth ed.). Canada: South-Western, Cengage Learning.

Appendix: Data analysis (SPSS)

A1: Results from T-tests; comparing the effect of each nudge against the control group of the month

Nudging using a reminder compared to the control group, February:

Group Statistics

	Group of test subjects	N	Mean	Std. Deviation	Std. Error Mean
Actual default	Reminder	500	.04	.186	.008
	Control	500	.11	.313	.014

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
										95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Actual default	Equal variances assumed	89.118	.000	-4.539	998	.000	-.074	.016	-.106	-.042
	Equal variances not assumed			-4.539	813.283	.000	-.074	.016	-.106	-.042

Nudging using a reminder compared to the control group, March:

Group Statistics

	Group of test subjects, February	N	Mean	Std. Deviation	Std. Error Mean
Actual default, March	Reminder	500	.07	.259	.012
	Control	500	.13	.332	.015

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
										95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Actual default, March	Equal variances assumed	33.857	.000	-2.868	998	.004	-.054	.019	-.091	-.017
	Equal variances not assumed			-2.868	941.591	.004	-.054	.019	-.091	-.017

Nudging based on the anchoring-effect compared to the control group, February:

Group Statistics

	Group of test subjects	N	Mean	Std. Deviation	Std. Error Mean
Actual default	Anchoring	500	.06	.230	.010
	Control	500	.11	.313	.014

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
										95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Actual default	Equal variances assumed	39.985	.000	-3.107	998	.002	-.054	.017	-.088	-.020
	Equal variances not assumed			-3.107	916.241	.002	-.054	.017	-.088	-.020

Nudging based on the anchoring-effect compared to the control group, March:

Group Statistics

	Group of test subjects, February	N	Mean	Std. Deviation	Std. Error Mean
Actual default, March	Anchoring	500	.13	.337	.015
	Control	500	.13	.332	.015

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
										95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Actual default, March	Equal variances assumed	.143	.705	.189	998	.850	.004	.021	-.038	.046
	Equal variances not assumed			.189	997.823	.850	.004	.021	-.038	.046

Nudging based on loss aversion compared to the control group, February:

Group Statistics

	Group of test subjects	N	Mean	Std. Deviation	Std. Error Mean
Actual default	Loss aversion	500	.07	.255	.011
	Control	500	.11	.313	.014

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Actual default	Equal variances assumed	19.937	.000	-2.213	998	.027	-.040	.018	-.075	-.005
	Equal variances not assumed			-2.213	959.161	.027	-.040	.018	-.075	-.005

Nudging based on loss aversion compared to the control group, March:

Group Statistics

	Group of test subjects, February	N	Mean	Std. Deviation	Std. Error Mean
Actual default, March	Loss aversion	500	.10	.303	.014
	Control	500	.13	.332	.015

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Actual default, March	Equal variances assumed	5.727	.017	-1.194	998	.233	-.024	.020	-.063	.015
	Equal variances not assumed			-1.194	989.651	.233	-.024	.020	-.063	.015

Nudging based on social norms compared to the control group, February:

Group Statistics

	Group of test subjects	N	Mean	Std. Deviation	Std. Error Mean
Actual default	Social	500	.13	.334	.015
	Control	500	.11	.313	.014

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Actual default	Equal variances assumed	3.095	.079	.878	998	.380	.018	.020	-.022	.058
	Equal variances not assumed			.878	993.741	.380	.018	.020	-.022	.058

Nudging based on social norms compared to the control group, March:

Group Statistics

	Group of test subjects, February	N	Mean	Std. Deviation	Std. Error Mean
Actual default, March	Social	500	.14	.349	.016
	Control	500	.13	.332	.015

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Actual default, March	Equal variances assumed	2.207	.138	.742	998	.458	.016	.022	-.026	.058
	Equal variances not assumed			.742	995.462	.458	.016	.022	-.026	.058

A2: Results from T-tests; comparing the effect of each nudge between February and March

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Reminder February	,04	500	,186	,008
	Reminder March	,07	500	,259	,012
Pair 2	Anchoring February	,06	500	,230	,010
	Anchoring March	,13	500	,337	,015
Pair 3	Loss aversion February	,07	500	,255	,011
	Los aversion March	,10	500	,303	,014
Pair 4	Social february	,13	500	,334	,015
	Social March	,14	500	,349	,016

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Reminder February & Reminder March	500	,694	,000
Pair 2	Anchoring February & Anchoring March	500	,630	,000
Pair 3	Loss aversion February & Los aversion March	500	,814	,000
Pair 4	Social february & Social March	500	,942	,000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	Reminder February - Reminder March	-,036	,186	,008	-,052	-,020	-4,317	499	,000
Pair 2	Anchoring February - Anchoring March	-,074	,262	,012	-,097	-,051	-6,315	499	,000
Pair 3	Loss aversion February - Los aversion March	-,032	,176	,008	-,047	-,017	-4,062	499	,000
Pair 4	Social february - Social March	-,014	,118	,005	-,024	-,004	-2,662	499	,008