Behavioral Economics as a Marketing and Communication Tool

NANNA NYGAARD & ALEXANDRA MARIA MUNK
STUDY NUMBERS: 34424 & 46657

MASTER'S THESIS:
MA IBC INTERCULTURAL MARKETING & CAND.MERC.(KOM)
SUPERVISOR: FLORIAN KOCK

CHARACTERS: 231.054
PAGES: 113

MAY 15 2018

CBS
COPENHAGEN
BUSINESS SCHOOL

ABSTRACT

This research contributes to the general understanding of nudging as an effective marketing and communication tool. Today, nudging has become a popular theme within most marketing and communication departments, and there seems to be a general conviction that nudging can be used for all kinds of issues with the intention of changing behavior of both employees, consumers and oneself. The purpose of this thesis is to broaden the understanding of when and how nudging can and cannot be used in a marketing and communication context. We challenge the conviction that nudging should be used uncritically to solve any kind of marketing and communication problem. Further, the research provides a better understanding of how certain heuristics function to influence behavior. Finally, we seek to determine whether the effect of an individual heuristic can be isolated to be able to identify and measure the actual outcome of a nudge.

To provide a valid answer for the problem that we attend to, we develop an experiment of our own, within which we seek to investigate the different heuristics and biases that come into function during a real-life nudge. Herein, we give special attention to the availability heuristic with the purpose of determining when a nudge must happen for it to have a decisive impact on behavior. To do so, we include two experimental units to which we apply the same treatment, i.e. the nudge, at two different points in time to be able to determine whether the bias of retrievability proves to impact the two groups differently. In this, we seek to isolate the effects of the availability heuristic from other heuristics and biases that prove to influence the behavior that we attend to.

The primary finding of the thesis is that the effects of the availability heuristic cannot be isolated from other heuristics and biases, as they prove to be interdependent. We find that certain heuristics may come into effect at one point of the decision-making process, whereas others influence other parts of the process, and others again increase or decrease the effects of the other biases and heuristics. Though, we did observe immediate differences between the two experimental units, and we argue that the individual heuristics might impact the experimental units differently due to the different times of treatment.

The findings lead us to a general discussion about the dynamics of heuristics in decision-making and the challenges that this poses to marketing and communication professionals.

RESUMÉ

Dette speciale bidrager til den generelle forståelse for nudging som et effektivt marketing- og kommunikationsværktøj. Nudging er i dag blevet et populært emne i de fleste marketing- og kommunikationsafdelinger, og der synes at være en general overbevisning om, at nudging kan bruges inden for en bred vifte af områder til at ændre adfærden hos både medarbejdere, kunder og en selv. Formålet med dette speciale er at udvide forståelsen for, hvornår og hvordan nudging kan og ikke kan anvendes i en marketing- og kommunikationskontekst. Heri udfordrer vi den generelle forståelse om, at nudging ukritisk skal anvendes til at løse ethvert marketing- og kommunikationsproblem. Samtidig giver specialet en dybere forståelse for, hvordan udvalgte heuristics kan bruges til at ændre adfærd. Endelig forsøger vi med specialet at klarlægge, hvorvidt effekten af et enkelt heuristic kan isoleres, med formålet at bedre kunne identificere og måle på det reelle udbytte af et nudge.

For at kunne give svar på problemformuleringen udvikler vi vores eget eksperiment, hvori vi vil undersøge de forskellige heuristics og biases, som kommer i spil under et virkeligt nudge. Vi sætter særligt fokus på availability heuristic med formålet at kunne bestemme, hvornår et nudge skal ske, for at det har en udslagsgivende effekt på adfærden. For at kunne bestemme det, gør vi brug af to forskellige eksperimentgrupper, som vi behandler på samme måde med et nudge, men på forskellige tidspunkter af dagen. Det gør vi for at kunne identificere, hvorvidt retrievability bias viser sig at påvirke de to grupper forskelligt. Hermed forsøger vi at isolere effekten af availability heuristic fra andre heuristics og biases, som kan have en indflydelse på den adfærd, som vi undersøger.

Den overordnede konklusion på problemformulering er, at virkningen af availability heuristic ikke kan isoleres fra andre heuristics og biases, da de alle viser sig at være gensidigt afhængige af hinanden. Gennem eksperimentet konstaterer vi, at bestemte heuristics kommer i spil på et givent tidspunkt i beslutningsprocessen, mens andre påvirker andre dele af processen, og andre igen styrker eller svækker virkningen af de andre biases og heuristics. Vi observerede dog en række umiddelbare forskelle mellem de to eksperimentgrupper, og vi påstår dermed, at de individuelle heuristics kan påvirke de to grupper forskelligt på baggrund af de forskellige tidspunkter, hvor nudget indtræffer.

Resultatet af analysen leder os over i en generel diskussion af de dynamikker, som vi observerer inden for heuristics i beslutningsprocesser og de udfordringer, som det skaber for marketing- og kommunikationsprofessionelle.

TABLE OF CONTENTS

ABSTRACT 2

RESUMÉ 4

INTRODUCTION 8
- Research Question 10
- Personal Motivation for the Thesis 11
- Delimitation 12
- Thesis Structure 13

LITERATURE REVIEW 15
- Rationality in decision making 16
- Dual Process Theory 17
- Heuristics and biases 23
- Nudging 29

HYPOTHESIS DEVELOPMENT 38

METHODOLOGY 44
- Philosophy of Science 44
- Research Design 46
- Experimental Design 47
  - Developing the experimental design 47
  - Validity and reliability 47
  - Experimental setup 48
  - Data collection 52
  - Limitations of the experiment 56
- Measurements 58

DATA ANALYSIS 62
- Expected Consumption 62
- Actual Consumption 67
- Expected Annual Consumption Ratio (EACR) 74
- | EACR | 78
- Hypothesis Testing 81

DISCUSSION **86**
Post-talk Analysis **86**
Discussion of findings **92**
Isolation of the Heuristics and Biases **99**
Can you Nudge a Habit? **103**
A Paternalistic Nudge? **105**
Nudging is not be for the Masses **107**

FUTURE RESEARCH **109**

CONCLUSION **111**

REFERENCES **115**

APPENDIXES **117**
Appendix 1: The five plates of fruit and vegetables 117
Appendix 2: Data observations 118
Appendix 3: The Y-Plate (from the Danish Veterinary & Food Administration) 127
Appendix 4: SPSS-values 128
Appendix 5: Descriptive statistics 134
Appendix 6: Independent t-tests 135
Appendix 7: Graphs of distribution 137

INTRODUCTION

"Nudge Yourself to Become Healthier: 4 Rules to Remember"
"Workplace Nudging Persuades People to Desirable Behavior"
"The Little Nudge that Makes a Big Difference to Student Grades"

Such headlines appear from a simple internet search on '*nudging'* and this rather simple tool seems to have become the answer to all kinds of issues, including workplace problems, self-enhancement, and environmental issues. With 5 simple rules you can nudge yourself to become more productive, you can nudge your employees to become a better workforce and create a better work environment and you can nudge your customers to buy more of your products.

Nudging has become a topic that every marketing and communication professional talks about and has an opinion on. It has proved to be an effective tool if used properly, and it has been successfully applied by governments as a complement to legislation, rules and bans. Traditionally, policy tools that have been used to influence people's behavior comprise only legislations, regulations and bans. More recently, insights from behavioral economics have suggested nudging as a new powerful tool that more successfully influences people's behavior and decision-making, as the theory is based on knowledge about human irrationality and the functioning of different systems in the brain in decision-making processes (Vlaev et al., 2016, p. 552).

A nudge is defined as "*any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives. To count as a mere nudge, the intervention must be easy and cheap to avoid. Nudges are not mandates. Putting fruit at eye level counts as a nudge. Banning junk food does not"* (Thaler & Sunstein, 2009, p. 6). From this rather simple definition, nudging seems to have become a tool that everyone can learn to apply. You can sign up for a nudging class and within just a few days, you will learn how to nudge. Such classes will teach you how to become a Choice Architect, as it is termed within this field of study. Nudging experts and nudging consultancies such as 'The Nudging Company', 'Kl.7', and 'Operate' have started to appear, offering nudging as a service to solve all kinds of issues.

Additionally, nudging and techniques for nudging is found in mainstream literature and publications about marketing and communication. The most recent and publicly commented piece of literature in a Danish context is Morten Münster's book: '*Jytte fra Marketing er desværre gået for i dag*', which proposes a 4 step framework on how to nudge: (1) Define the desired behavior, (2) Make an analysis of the barriers, (3) Design your solution, and (4) Test your solution (Münster, 2017). The book functions as a practitioners guide on how to create nudges, and communication professionals within several areas have used it widely. Evidently, nudging has become common property and everyone can proclaim himself or herself to be a Choice Architect.

However, one might question whether nudging is in fact so simple and whether everyone should have the opportunity to become a Choice Architect. The Choice Architect will naturally have the power to influence other people's choices and decision-making and with such power follows a responsibility. Therefore, we raise concerns about how-to guides and nudging classes and question how they might negatively impact the power of nudging.

Theories on nudging have evolved with its popularity, especially over the past 5 years. Introduced by Thaler and Sunstein (2008), nudging started out as a way to improve choices by exploiting the knowledge of irrational decision-making of human beings. The irrationality is caused by cognitive limitations of the human brain, which results in that certain heuristics impact the decision-making situation, i.e. the decision is biased by heuristics. The first known examples of nudging exploit these heuristics and biases and use them to impact decision-making. Today, nudging has expanded to involve all kinds of interventions and elements that can impact decision-making, not only being bounded by heuristics and biases. In theory, nudging can be anything as long as it influences people's behavior to the better with the help of interventions, which in theory should not have an effect but do in practice. Such thinking suggests that a difference between econs (Homo economicus) and humans (Homo sapiens) exists.

From this development of the nudging theory, we raise concerns about how one can determine if a nudge has a provable effect if the underlying causes are of various forms and ideas with a minimum theoretical background. Problems of determining exactly what works

and why, might arise. This further questions the degree to which nudging can be used as an effective marketing and communication tool.

Research Question

This research sets forth to investigate how nudging can indeed be an effective marketing and communication tool and seeks to determine the uncertainties that one should be aware of when using nudging in a marketing context.

The research gives attention to the use of a certain heuristic as a means to nudge. More specifically, we look into the availability heuristic and investigate how most recent events and easily recollected memories impact decision making in everyday situations. The research seeks to investigate how the availability heuristic causes biased decision-making illustrated by an experiment that centers on consumption. The experiment is conducted in a canteen. Such a setup is often seen within the theories of nudging.

The overall question that guides the research is:

> *How can the availability heuristic be used as a nudge to impact decision-making, and how can one isolate and measure the effects of a specific heuristic? Further, what are the consequences of the uncritical and general use of nudging in mass society?*

To provide an answer for *how* the heuristic can be applied, we will make the nudging experiment of our own. The experiment seeks to investigate how the availability heuristic can be used as a means to nudge people's consumption behavior, in the same sense that other heuristics have been used to impact decisions on health and well-being (Thaler and Sunstein, 2009, p. 7). The experiment is developed from the knowledge about nudging that we have obtained from thorough investigation of the field and reviews of previous research. Such research will be presented in the literature review.

When investigating the effects of the availability heuristic, we seek to get a better understanding of exactly when a nudge has to take place in order for it to have the most optimal

effect on one's behavior. The experiment sets forth to test when an incident or exposure to a message must happen for it to have a decisive effect on behavior, i.e. to what extent the memory is available for recollection and thus impacts decision making. This thorough understanding of the availability heuristic and the effects of recollected memories expand the already accessible research within this field. This can thereby help companies and Choice Architects understand how nudging by the availability heuristic can best be applied as a managerial tool.

In seeking to answer *how* one can isolate the availability heuristic and measure the effects of the nudge, we develop a range of hypotheses that guide the experiment and the following data analysis. The analysis will function as a means to measure whether the nudge succeeds in impacting the behavior of the participants of the experiment according to our expectations. By developing a range of hypotheses that propose different outcomes of the experiment, we ensure that the analysis covers several aspects that might be the cause of a potential change in behavior.

Finally, the discussion will help provide an answer for the final part of the research question and highlight the potential consequences of the general use of nudging. The analysis will lead to this discussion of how nudging should be applied to be an effective marketing and communication tool. This involves considerations about what marketers should be aware of when developing nudges and it comments on the potential risks of every communication professional proclaiming themselves to be Choice Architects.

Personal Motivation for the Thesis

Already at the initial phase of the thesis, we found the topic of nudging to be very interesting. The topic came to our attention, as well as it came to most marketing and communication professionals' attention, during 2016/17. Following this period, the topic was slightly hyped and it suffered from being a buzzword within marketing during the last half of 2017. At this point we still believed that nudging, if used properly, can be a very efficient marketing and communication tool and that it can have a decisive impact on behavior that cannot be obtained by e.g. rules and regulations. Therefore, we set forth to investigate this field

from a critical perspective, seeking to broaden our understanding of how nudging can be used as an effective marketing and communication tool.

It quickly came to our attention that nudging is more complex than first understood. This only strengthened our critical thoughts on the application of nudging and it led us to question whether nudging should really be a tool to be used by the masses. With this thesis, we seek to open up the discussion about when and why nudging is an appropriate tool and we try to spark a general critical debate amongst marketing and communication employees. In this, we hope to provide marketers with an ability to critically respond to a potential push for nudging from marketing management or top management as a result of the general hype around the topic. This is to avoid that marketers throw themselves into a nudging campaign without the proper knowledge, which in the end fails to succeed. Further, this should help ensure that nudging as a tool for marketing and communication is not diluted and misused, which might very well be the case at the moment. Therefore, this thesis seeks to discuss whether nudging should be applied by all and everyone and it proposes that nudging might not be the proper solution to all kinds of issues and for all kinds of communication problems.

Delimitation

When discussing the concept of nudging, it is often associated with the concepts of priming and framing. Framing is defined as "*setting an approach or query within an appropriate context to achieve a desired result or elicit a precise answer*" (Business Dictionary, 2018a), and priming is defined as: *"the activation of various mental construct unbeknownst to individuals via perception of external stimuli*" (Weingarten et al., 2016, p. 472). In this thesis, we do not give attention to the concepts of priming and framing, but we give the full attention to the element of nudging, though acknowledging that priming, framing and nudging are elements that often times converge. If applied in a combination, nudging, priming and framing might be even more powerful tools to influence the behavior of people but in this research, we seek to investigate the concept of nudging in isolation.

This further means that within our experiment, we do not want to achieve a desired result within a framed context, which is set to evoke a specific answer. We want the answers

from the participants to be as truthful and frank as possible. Hence, we do not frame. Second, we are in general not interested in the topic of health and healthy eating but we use health as an illustration for how to use nudging as a marketing and communication tool. One might argue that we prime the participants in the direction of healthy consumption because when the participants are exposed to the message of the experiment concerning consumption, they could be inclined to think that the subject of the experiment concerns health. However, this is not the scope of the research and therefore we do not seek to prime the participants.

Through these definitions of the concepts, we seek to make it clear that we delimit our focus to *nudging* throughout this thesis. We seek to exclude priming and framing as much as possible from the experiment to ensure that we can isolate the effects of the nudge.

Thesis Structure

With this initial chapter we set the scene of the thesis by introducing the field of research and the general problem that we will attend to in this research. A range of questions arises from the introductory part, which forms the basis for the overall research question of this thesis. The chapter is followed by a literature review in which we critically review and present the existing theories within the field of investigation. In this, we comment on the validity of the chosen literature on which the research is based. From the findings in the literature review, a set of hypotheses is developed. The hypotheses help guide the development of the experiment and the following data analysis. Hereafter, the methodology chapter gives an account for the methodical considerations and choices made prior to the experiment. This chapter is divided into four sections, including philosophy of science, experimental design and the data collection, which together form the basis of the thesis' overall research design. In the analysis chapter, we analyze the data from the experiment. The result of the analysis enables a test of the different hypotheses where we either accept or reject the different hypotheses. This leads us towards an answer for the overall research question. Hereafter, observations and questions, which have emerged during the experiment, form a basis for the discussion together with the general topics that we have set forth to investigate. From this, we present recommendations for future research within the field of nudging, and we point to new research directions and perspectives, which have

emerged through this research. The final chapter of the thesis concludes on the findings of both the analysis and the discussion and provides an answer for the research question.

LITERATURE REVIEW

This chapter examines relevant literature on the topics of decision-making and behavioral economics to which the theory of nudging belongs. It comes around theories of heuristics and biases that impact decision making, as well as theories on nudging and libertarian paternalism and it provides a critical account for the current state of decision-making theory. The review examines both major writings, e.g. from Nobel Prize winners Daniel Kahneman and Richard Thaler, as well as more recent research on the topic. The review provides an understanding of how this piece of research enters into the general theories of nudging and decisions making, and how it sets forth to contribute to the current understanding of the topic within the fields of marketing and communication. Finally, it enables the development of key hypotheses that will guide the research.

The topic that guides this research is the theory of *Behavioral Economics* and more specifically *Nudging*. Behavioral economics is the science of human decision-making and the study of how humans make decision in the real world (Vlaev et al., 2016, p. 551). Behavioral economics unifies insights from psychology and economy and provides new ways of thinking about the drivers of certain kinds of behavior and on the barriers to behavior (Vlaev et al., 2016, p. 551). Nudging is part of the Behavioral Economics-thinking and it concerns the way with which humans are psychologically impacted in decision-making situations. Throughout this review we give attention to theories and topics related to the nudging literature to obtain a thorough understanding of nudging and the theories that underlie.

The chapter starts with a short review of general thinking of decision-making and rationality, which is found relevant for the general understanding of the nudging and behavioral economics theories. Hereafter follows a review of theories about human cognition and systems in the brain, which are fundamental principles in the understanding of decision-making processes. This leads to an overview of heuristics, biases and other factors that are said to impact decision-making. Finally, we introduce the first thinking of nudging, termed at first as Libertarian Paternalism, and we present as well some more recent ideas and elements within the theories of nudging.

The review includes sources from highly ranked journals and researchers as well as mainstream literature, e.g. Kahneman's award winning book '*Thinking fast and slow*' and Thaler & Sunstein's book '*Nudge*'. Though, as these books are written by Nobel Prize winners and are based on highly ranked literature, the sources and findings herein are considered to be valid. In general, we strive to use literature that is published in journals, which are ranked 3 or 4 in the ABS Academic Journal Guide from 2015. The rank indicates a journal's impact and quality based on peer reviews, editorial and expert judgments as well as statistical information about the number of citations (Chartered Association of Business Schools, 2015). Furthermore, the journal rank indicates the place of the journal within its field, the difficulty of being published and prestige identified with the journal (Chartered Association of Business Schools, 2015). This provides us with the most valid and the most acknowledged literature within the field, on which we base our research.

Rationality in decision making

Decision-making is defined as *"the thought process of selecting a logical choice from the available options… (where) a person must weigh the positives and negatives of each option… forecast the outcome of each option as well, and… determine which option is best for that particular situation"*. (BusinessDictionary, 2018b) In reality, almost no decisions are made on these bases as human beings make several hundred decisions every day. Most are made unconsciously while others are made consciously, and even conscious decisions are often based on a "good-enough" approach. In most situations, we have neither the necessary information available nor the cognitive ability to make logical decisions that consider all possible options.

Much investigation has gone into the field of decision-making. What seems to be a recurring theme of discussion is the question of rationality versus irrationality in decision-making processes. Today, most researchers acknowledge that decision-making processes do not follow the principles of Homo economicus. Such thinking goes against the definition of decision-making as proposed above, and it is still a fairly new way of thinking about decision-making. In this, more researchers propose that human decision-making is somewhat irrational, or at least *bounded rational*. Bounded Rationality is a recurring term within much literature on decision-making, e.g. commented by Richard Thaler and Herbert Simon. Ac-

cording to Simon, bounded rationality is a redefined understanding of rationality that takes the cognitive limitations of the decision maker into consideration (Simon, 1990, p. 15). Herein, Simon proposes that human decision-making is rational but rational in the sense that the human being makes the best possible decision given the information that was available to him or her at the time of the decision-making. Such ideas about bounded rationality do, according to Simon, deviate from the subjective expected utility (SEU-theory) that underlies neoclassic thinking of profit maximization. Simon suggests that one must accept estimations and uncertainties with regards to probability in decision-making, and as well realize that most people settle with a good-enough alternative (Simon, 1990, p. 15). This is due to the limited cognitive capabilities of the human brain. However, even on such terms, human decision-making is not irrational, Simon argues. Instead, it is bounded rational.

This thinking underlines the relevance of this research and legitimates the problem area that we attend to. While agreeing with Simon in his observations of the limited cognition of human beings and its implications for rationality in decision making, we move forward with a look into the cognitive elements, which are at play during decision-making processes. Herein, we look into theories on dual processes and cognitive systems as well as heuristics and biases that influence decision making.

Dual Process Theory

Theories on dual processes have been part of cognitive and social psychology studies for more than 40 years. Throughout the years, many authors have shown interest in dual processes and their effects (Evans, 2008, p. 256), which have resulted in different ideas and perspectives on the matter. One element that unites the different theories is the idea that different ways of processing information exist. In the literature, these ways of processing have been termed in several ways. Some of the most well known terms that group these different ways of thinking are the *'System 1'* and *'System 2'* by Kahneman and *'The Automatic System'* and *'The Reflective System'* by Thaler and Sunstein. Others claim that there might be more than two systems in the brain. Although the idea of dual process theories have existed for many years, it is only within the past 20 years that the idea of System 1 and System 2 have come to be commonly used by psychologists and the general masses

(Evans, 2008, p. 270). However, there seems to be conflicting ideas about the two ways of thinking and how they are linked to systems in the human brain. In the following, four theories are presented, the first two being Thaler & Sunstein's Automatic and Reflective system and Kahneman's System 1 and System 2. Thaler, Sunstein and Kahneman seem to agree to a large extent on the characteristics of two separate cognitive systems. Second, we present two recent perspectives on this system thinking. First, we look at Jonathan Evans' critique of the simplistic presentation of the systems in the human brain. Evans challenges the general literature on dual processes and its superficial treatment of the complex systems in the brain. Hereafter, we look to a very recent piece of literature by Ivo Vlaev and Paul Dolan. Vlaev and Dolan propose that at least three systems can be found in the brain and that they all impact actions and decisions, and having knowledge of these three systems can help change behavior. That is why Vlaev and Dolan present a theory that they call Action Change Theory, which utilizes these three systems in the brain. Their research builds on very recent findings within the field of cognitive science, which uncovers how human beings acquire complex behaviors.

The Automatic System and The Reflective System by Thaler & Sunstein

Richard Thaler is one of the world's leading behavioral economist and Cass Sunstein is a pioneer and leading scholar in Behavioral Law (Hansen, 2016, p. 6). Thaler and Sunstein have worked together on a range of different theories and ideas, and they have dedicated most of their lives to behavioral economics. They are probably most famous for their best-selling book '*Nudge*', which will be further introduced at a later stage. In this book, they dedicate the whole introduction to a summary of their findings from four decades of behavioral economics.

As mentioned, Thaler and Sunstein propose that the human brain has two ways of thinking; one that is intuitive and automatic and one that is reflective and rational (Thaler & Sunstein, 2009, p. 19). The automatic system is characterized by being fast, uncontrolled, and unconscious, whereas the reflective system is characterized as slow, controlled, and self-aware or conscious. To illustrate the difference between the two systems, they use the following example: *"The automatic system says that "the airplane is shaking, I'm going to die", while the reflective system responds, "Planes are very safe""* (Thaler & Sunstein,

2009, p. 21). Sunstein and Thaler argue that the automatic system to a large extent can be described as one's "*gut feeling*". The gut feeling often relies too heavily on intuition and sometimes this results in mistakes (Thaler & Sunstein, 2009, p. 21). However, decisions based on the automatic system can often times also result in the right choice. According to Thaler and Sunstein, there is no right or wrong between the different systems and their ability to make decisions.

System 1 and System 2 by Kahneman

Daniel Kahneman has been highly involved in theories of dual processes. Kahneman has conducted much research within the field of psychology, often times in collaboration with Amos Tversky, and dual processes have been one of many elements under study. Kahneman's most popular writing about dual processes is his bestselling book '*Thinking Fast and Slow*', dedicated to the ideas of the two ways of processing information. Kahneman terms this System 1 and System 2 (Kahneman, 2011, p. 20). *"System 1 operates automatically and quickly, with little or no effort and no sense of voluntary control"* and *"System 2 allocates attention to the effortful mental activities that demand it, including complex computations. The operations of System 2 are often associated with the subjective experience of agency, choice and concentration"* (Kahneman, 2011, p. 21). System 1 is similar to what Thaler and Sunstein term the automatic system, whereas System 2 is similar to the reflective system that is activated at complex tasks. In his book, '*Thinking Fast and Slow'* Kahneman discusses some of the instances where System 1 and System 2 individually come into play, and it becomes evident that each system has its own strengths and weaknesses. Kahneman describes the systems as *"agents with their individual abilities, limitations and functions"* (Kahneman, 2011, p. 21). Evidently, Kahneman agrees with Thaler and Sunstein in the notion that there is no right or wrong system, or system thinking, but that both ways of processing information have strengths and weaknesses depending on the choice situation.

Both systems are active when we are awake, though System 2 is mostly running at a minimum level. System 2 constantly receives information from System 1. When System 1 runs into a problem that it cannot solve, it calls for System 2. This collaboration is fine at most times. However, one can at times run into difficulties with the System 1-thinking. System 1

has its limitations, which are caused by cognitive heuristics. These can result in biased decision-making and such biases might result in systematic errors in the decision-making (Kahneman, 2011, p. 25). This is very much in line with Thaler and Sunstein's theories about dual processes and decision-making. The heuristics and how they can impact the decision-making are presented in detail at a later stage.

Jonathan Evans' critique of dual-process theory and dual systems

Now, we look into some of the critiques of the dual-process theory. As mentioned, more researchers have argued that the dual processes are too simplistic. One of those is Jonathan Evans. Evans has contributed to the general theories of cognitive abilities and biased decision-making from a rather critical account. In this, he has sought to broaden the understanding of dual-process theories and decision-making. Evans argues: *"it is most certainly wrong to think of System 1 as one system, all of which is old and shared with other animals. Equally, it is probably a mistake to think of System 2 as the conscious mind, all of whose processes are slow and sequential"* (Evans, 2008, p. 271). Instead, Evans considers the possibility that one system operates only with System 1 processes, the intuitive and fast thinking, whereas a second system might work with a mixture of System 1 and System 2 processing. His argument for such a claim is that System 1 processes can evolve over time, as they are simply any process in the mind that can work automatically without the use of memory space (Evans, 2008, p. 271). Such processes can be developed from System 2 to System 1 thinking with training and experience, e.g. when learning a new language. In conclusion, Evans seems to agree that there are indeed different ways of processing information, though how we process certain types of information can evolve over time, but he discards the idea that such information processing is bound in two single operating systems in the brain.

Action Change Theory by Vlaev & Dolan

Ivo Vlaev and Paul Dolan seem to agree with Evans' critique of the too simplistic understanding about processes and systems in the brain. Vlaev and Dolan are both professors of Behavioral Science and have done much research within the field of behavioral economics and nudging. Most recently, they published a paper focusing on nudging as a tool to improve behaviors around health. With their own theory on how to impact behavior, Ac-

tion Change Theory, they present the argument that dual-process thinking is too simplistic and that more cognitive systems exist in the brain (Vlaev & Dolan, 2015, p. 71). Vlaev and Dolan propose that there are three core brain systems that guide decision-making; the goal-directed, the impulsive and the habitual system (Vlaev & Dolan, 2015, p. 71). These brain systems can be independently or jointly activated to achieve behavior change. Vlaev and Dolan's research is backed by recent findings within cognitive neurosciences, which investigate how the human brain responds and works under complex decision-making (Vlaev & Dolan, 2015, p. 71).

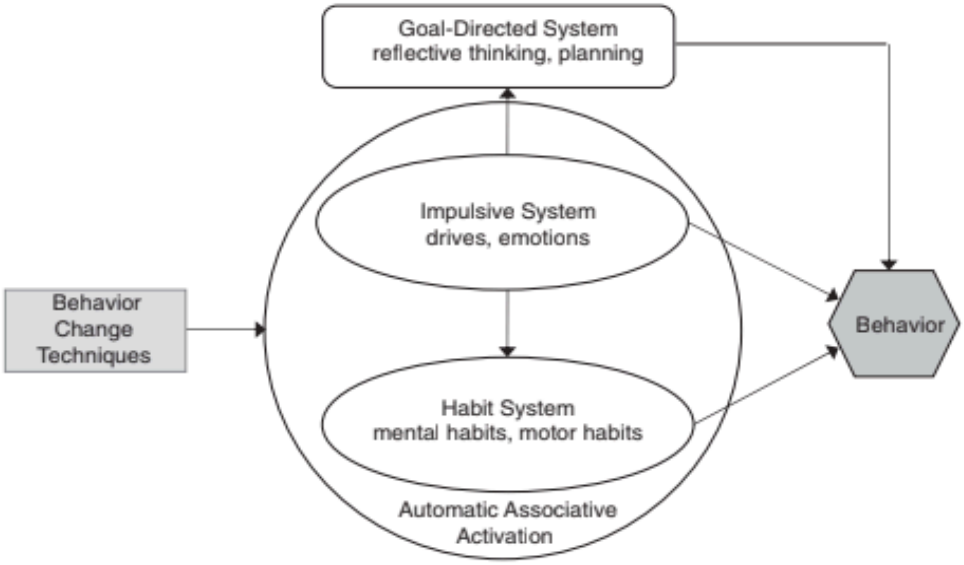


Figure 1: Self-Regulatory Processes Involved in Behavioral Change. Source: Vlaev et al. (2016).

If compared to previous thinking about the brain systems, *the Goal-Directed System* is similar to System 2 thinking, whereas both the *Impulsive* and the *Habit System* belong to System 1. The goal-directed system is defined by *"model-based reasoning to calculate action-outcome contingencies"* (Vlaev et al., 2016, p. 552). This involves reflective thinking and calculations of probability, which is characteristically for System 2 thinking. The habit system is centered on "*learning through repeated practice in a stable environment to assign values to a variety of actions proportionally to the rewards and punishments received as a result of executing those actions*" (Vlaev et al., 2016, p. 552). The habit system is divided into *mental habits* and *motor habits*. Mental habits are defined as automatic processes where the automaticity is achieved through frequent execution of cues by the use of connections in the long-term memory (Vlaev & Dolan, 2015, p. 74). Motor habits are de-

veloped through repetition as well-practiced actions are repeated in the same circumstances with the same states or stimuli being present (Vlaev & Dolan, 2015, p. 74). Through this process, the motor habits are automatically triggered by environmental cues and thereby easier to perform again and again. These habits are usually a result of a desired goal, but once one have acquired these habits, they can be performed without the goal in mind (Vlaev & Dolan, 2015, p. 74). Finally, the impulsive system triggers innate behavior, i.e. evolutionary responses to certain stimuli. Such stimuli trigger automatic behaviors that lie true to the human nature (Vlaev et al., 2016, p. 253).

To sum up, all the above researchers agree that there are different ways of processing information at different times and situations. Some situations require only intuitive information processing, which leads to fast and often times unconscious decisions, whereas in other situations the mind will call upon reflective thinking and reasoning to come to a conclusion. The mind has limited cognitive abilities and cannot engage in reflective thinking all the time. Therefore, intuitive decisions are a necessity regardless of whether such thinking is thought of as belonging to a certain system in the brain or whether they are just one of several ways of thinking and making decisions. The problem becomes that such intuitive thinking is limited and can at times result in mistakes and flaws. This happens as people make decisions based on what Thaler & Sunstein call '*Rules of thumb*' and Tversky & Kahneman call '*heuristics*' (Thaler & Sunstein, 2009, p. 22). Amos Tversky & Daniel Kahneman were the first to develop this way of thinking about biased decision-making. According to them, heuristic judgment can result in biased reasoning. Not saying that biased thinking is always negative, though, they recognize that biases can be beneficial at times but at other times faulty, and in such cases, they are something to overcome. Dolan and Vlaev elaborate on this thinking in their Action Change Theory and propose a multiple of bases, not only limited to heuristics, on which one can change behavior based on the cognitive abilities or disabilities of the human brain. This way of changing or improving behavior is what will later presented as nudging.

A thorough understanding of these brain systems and their functionalities strengthen the understanding of how one with the help of nudges can improve or impact judgment and

the decision-making processes. Therefore, research of heuristics and nudging by interventions will be presented in detail below.

Throughout this thesis, we apply the concepts of System 1 and System 2, as Kahneman and Tversky named them. These terms are most commonly used in the nudging literature, which is why we find it most appropriate to also use these terms. However, we are aware that this division into only two systems is probably too simplistic and that one should pay attention to when and how each of the different systems are activated and how System 1-thinking might evolve over time.

Heuristics and biases

Heuristic judgment is caused by a limited validity of the human perception. It is concerned with the fact that the resources of the human brain are limited and that often time humans rely on heuristic rules when making complex decisions or judgments. Heuristics are in their sense a natural element of human cognition. However, relying on such heuristic rules can in some cases result in mistakes or faulty judgments. This is called being biased. According to Tversky and Kahneman, biases are the negative outcome of applying heuristic rules that lead to faulty decisions or judgment (Tversky & Kahneman, 1974, p. 1124).

In order for people to complete complex tasks such as assessing the probability and predicting values of uncertain events, people rely on a limited number of heuristic principles (Tversky & Kahneman, 1974, p. 1124). This idea was presented by Tversky and Kahneman in 1974 in their leading article: '*Judgment under Uncertainty: Heuristics and Biases*'. The article describes the simplifying shortcuts of intuitive thinking by explaining three heuristics and 20 biases to which these heuristics can lead. They describe it as: "*Many decision are based on beliefs concerning the likelihood of uncertain events such as the outcome of an election, the guilt of a defendant, or the future value of the dollar. These beliefs are usually expressed in statements such as "I think that…," "chances are…," "it is unlikely that…,"and so forth*" (Tversky & Kahneman, 1974, p. 1124). The article links heuristics and judgment and demonstrates the role of heuristics in judgment and decision-making.

The first three heuristics to be introduced by Kahneman and Tversky are representativeness, availability, and adjustment and anchoring. According to Kahneman and Tversky, people rely on these heuristics in decision-making situations to be able to assess probability and to predict values. The representativeness heuristic is normally applied when being asked to judge the probability of an object or event belonging to a certain group or process. The availability heuristic comes into play when people have to "*assess the frequency or the probability of an event by the ease with which instances or occurrences can be brought to mind*" (Tversky & Kahneman, 1974, p. 1127). The last heuristic concerns adjustments from an anchor, which is normally used in numerical predictions where a base value, the anchor, becomes the reference point of judgment. The availability heuristic is of primary interest for this research but we will as well comment on other heuristics and biases, which can impact the decision-making that this research sets forth to investigate.

Availability heuristic

As more frequent events are usually easier to recall or more available than less frequent events, the availability heuristic often leads to inaccurate predictions of probability (Bornstein & Emler, 2000, p. 100). For example, if one is to assess the risk of cancer among old people, one will have a tendency to assess the probability based on occurrences among acquaintances. Reliance on availability can therefore lead to predictable biases. According to Tversky & Kahneman, four different biases exist as an element of the availability heuristic. The four biases are retrievability of instances, effectiveness of a search set, imaginability and illusory correlation (Tversky & Kahneman, 1974, p. 1127-1128). What is of interest to this research is primarily the bias caused by retrievability of instances and the bias of imaginability.

The ease with which one can retrieve an instance will impact decision-making. Several factors may have an impact on the retrievability of instances. Generally, instances that are more recent will be easier to retrieve than earlier instances. (Tversky & Kahneman, 1974, p. 1127). Such tendency to rely on recent instances can bias the decision-making.

Imaginability can as well bias the decision-making. If a certain outcome or situation is easy to imagine, it might be more available and hence impact the judgment of probability. However, the ease of imaginability does not always respond to reality. Kahneman and Tversky

present an example of this: "*The risk involved in an adventurous expedition… is evaluated by imagining contingencies with which the expedition is not equipped to cope. If many such difficulties are vividly portrayed, the expedition can be made to appear exceedingly dangerous, although the ease with which disasters are imagined need not reflect their actual likelihood"* (Tversky & Kahneman, 1974, p. 1128). This tendency can work the other way around as well, if the probability of risk is to complex to grasp and hence to imagine, the probability of the event to occur might be underestimated (Tversky & Kahneman, 1974, p. 1128).

Other heuristics and biases

Throughout the years, as much investigation has undergone the field of nudging, theories of heuristics and biases as well as other techniques to nudging have evolved in various ways. Many researchers each seek to contribute to the current understanding by defining a new set of heuristics and biases that in various ways impact decision-making. Both Vlaev & Dolan, Thaler & Sunstein and other researchers that we have come across, all offer their own version of a set of heuristics and biases that can influence behavior. Some of them are more or less the same under different names or with little variation, whereas others are a result of new ways of interpreting the concept of heuristics. As mentioned, the means to nudge involve more than just heuristics, and most of these researchers have as well developed other techniques that can be used to nudge. However, as this thesis focuses on nudging by heuristics we give the primary attention to this. In the following, we present four other heuristics/biases, which are found relevant for this research; the Anchoring and Adjustment heuristic, the Commitment heuristic, the Optimism bias and the Social Desirability bias.

Anchoring and Adjustment heuristic

"In many situations people make estimates by starting from an initial value that is adjusted to yield the final answer" (Tversky & Kahneman, 1974, p. 1128). Such an initial value is an anchor, which every estimate and judgment of probability is based upon. A person's level of knowledge about a certain subject forms the basis of ones anchor. By having the anchor, a person will adjust in accordance with new subjective experiences, learning and challenges and thereby form a new subjective anchor (Thaler & Sunstein, 2009, p. 23).

Different anchors result in different estimates and in this sense bias judgment and decision-making. For example, in negotiation one will often work with an anchor. The first bid for price or salary will become the anchor for all other bids. Inevitably, an anchor and the adjustment hereto bias decision-making.

Commitment heuristic

Commitments entail asking people to make a verbal commitment by stating to execute a specific action or to achieve a desired goal (Vlaev & Dolan, 2015, p. 85). After stating such an action or a goal, one will feel committed to achieve this. If the goal is achieved, it might result in a higher social status. If the goal is not achieved, one might risk losing status. This is the influence of the commitment heuristic (Vlaev & Dolan, 2015, p. 85). Commitments work to influence decision-making, as people after stating a goal publicly will seek to be consistent with their promise. A commitment could for example be a promise to go to the gym or to stop smoking. The power of the commitment is usually higher if the goal or action is written down (Vlaev & Dolan, 2015, p. 85).

Optimism bias

People tend to think of themselves as logical and rationally thinking human beings. However, researchers believe that the human brain is sometimes too optimistic for its own good (Cherry, 2017). This is known as the optimism bias. Unrealistic optimism is a common feature and do often appear in social contexts. People can be unrealistically optimistic both when the stakes are high and when the stakes are low (Thaler, Sunstein, 2009, p. 32). The bias leads one to believe that one will be less likely to suffer from adversity and more likely to achieve success than what is realistic to obtain. Examples of such instances are e.g. lotteries that are successful partly because of unrealistic optimism, or smokers who often exaggerate the statistical risks of smoking and who often believe that they are less likely to be diagnosed with diseases than nonsmokers (Thaler & Sunstein, 2009, p. 33).

Social Desirability bias

The Social Desirability bias is defined as people's tendency to present oneself in the best possible way (Fisher, 1993, p. 303). This becomes evident when people respond or act in

a manner that is perceived socially desirable and acceptable and viewed favorably by others (Fisher, 1993, p. 303). On one hand, the outcome of this can result in over-reporting good or desirable behavior, and on the other hand, it can result in underreporting bad or undesirable behavior (Fisher, 1993, p. 304). The social desirability bias is more prominent when dealing with sensitive subjects such as personal income, health, family and self-esteem.

Should heuristics and biases be overcome?

The concluding remark on Kahneman and Tversky's research on heuristics and biases state: *"A better understanding of these heuristics and of the biases to which they lead could improve judgements and decisions in situations of uncertainty"* (Tversky & Kahneman, 1974, p. 1131). From this comment, it seems that Kahneman and Tversky consider such biased decision-making as something to be overcome. This can be done by recognizing and informing oneself about the limitations of human cognition and how such limitations can distort judgment and decision-making.

However, all researchers do not share this view of heuristics and biased decision-making as something to overcome. One example is Gigerenzer, who instead proposes that heuristics and biases are a necessity for decision-making, and that it can at times even result in better decisions and judgments.

Gerd Gigerenzer has as well studied several topics within psychology. He is best known for his work on bounded rationality and heuristics. His theories about heuristics are of special interest to this work, as it challenges some of Kahneman and Tversky's ideas and their views on the value of biases and heuristics. Gigerenzer challenges specifically the idea that biased decision-making is faulty or negative and should be overcome.

Gigerenzer argues that decisions are made as the mind applies logic, statistics or heuristics (Gigerenzer et al., 2011, p. 452). At first, logic and statistics was perceived as the rational bases for decisions, whereas heuristics were found to be irrational and faulty, as already proposed. This understanding is mistaken according to Gigerenzer. He points to a range of researchers, who go against this line of thought and argue that heuristic reasoning might be just as accurate as logical and statistical reasoning. One of these researchers

is Simon with his theory of bounded rationality (Gigerenzer et al., 2011, p. 452). With this, Gigerenzer challenges the belief that heuristics should not be treated as an equal base for decision-making, and with his research he seeks to contribute to a deeper understanding of the element of heuristics in decision-making. (Gigerenzer et al., 2011, p. 453).

According to Gigerenzer, a heuristic is characterized as "*a strategy that ignores part of the information, with the goal of making decisions more quickly, frugally, and/or accurately than more complex methods*" (Gigerenzer et al., 2011, p. 454). Alone the wording of Gigerenzer's definition of a heuristics makes it clear that his understanding of heuristics and the biases connected here to, is of a more positive character, believing that heuristics can lead to accurate and fast decisions. He argues that *"the goal of making judgements more accurately by ignoring information is new. It goes beyond the classical assumption that a heuristic trades off some accuracy for less effort"* (Gigerenzer et al., 2011, p. 455). This underlines the claim that Gigerenzer is not of the conviction that heuristics and biases lead to error-prone information processing, and he further argues that heuristics can even be relied upon consciously (Gigerenzer et al., 2011, p. 455), i.e. in System 2. Therefore, according to Gigerenzer heuristics and biases are not limited to System 1 thinking, but can be used both unconsciously and consciously to make better decisions and judgments.

If attending to the idea of bounded rationality in combination with Gigerenzer's arguments on heuristic decision-making, one could propose that biased decision-making should not always be overcome. The cognitive limitations of the human brain mean that we, as human beings, are rarely able to make fully informed and rational decisions. In such cases, we must make decisions based on a limited rationality, which can often times result in the best possible outcome, even though the decision is biased by heuristics.

We do not seek to give a final answer to whether biased decision-making should be overcome but throughout the thesis we give attention to this problem area, as heuristics and biases are, to our conviction, valuable bases for nudging to succeed. This will be clear from the section below where nudging is presented.

Nudging

Nudging was developed by Thaler & Sunstein in 2003. They first introduced the theory under the term '*Libertarian Paternalism'*, which came to be developed into nudging, as we know of it today. This section introduces the theories of Libertarian Paternalism and the following theories of nudging. Today, the concept of nudging is well acknowledged and has been further developed by several authors and applied in various fields of study. In the following, we look into the first theory of nudging as well as more recent literature and research within the field.

Libertarian Paternalism by Thaler & Sunstein

Thaler & Sunstein introduced the concept that they name 'Libertarian Paternalism' in May 2003 in the paper '*Libertarian Paternalism*' (Thaler & Sunstein, 2003a). It is the concept that is later to be developed into nudging as most know of it today. Libertarianism is a political philosophy in which liberty is considered the primary political value. The philosophy is similar to the political philosophy of liberalism and the libertarians do as well advocate strongly for the right to liberty. They believe in the freedom of will, as long as those actions do not limit the freedom of others (Encyclopædia Britannica, 2018a). Paternalism is an attitude or a practice that is usually understood as *"an infringement on the personal freedom and autonomy of a person with a beneficent or protective intent"* (Encyclopædia Britannica, 2018b). At first, one might argue that libertarianism and paternalism are contradictory concepts that cannot be combined. However, Thaler and Sunstein argue that Libertarian Paternalism is not an oxymoron. This argument is presented only half a year after publishing the paper '*Libertarian Paternalism*' in May 2003. In their second paper '*Libertarian paternalism is not an oxymoron*' (Thaler & Sunstein, 2003b), published in the fall of 2003, Thaler and Sunstein argue that economists should rethink their views on paternalism and realize that paternalistic actions are inevitable and not always negative.

Thaler & Sunstein define paternalistic actions as actions *"with the goal of influencing the choices of affected parties in a way that will make those parties better off"* (Thaler & Sunstein, 2003a, p. 175). What is most important and probably also most difficult is to define what '*better off*' is, as this can vary between individuals and situations. To overcome this barrier, Thaler & Sunstein propose making a cost-benefit analysis that determines the full

ramifications of any design architecture (Thaler & Sunstein, 2003a, p. 178). However, this is not always possible. Therefore, Thaler & Sunstein propose instead three possible methods to determine whether such actions can be characterized as being paternalistic and in the best interest of the involved parties (Thaler & Sunstein, 2003a, p. 178). The three methods are (1) selecting the approach that the majority would select if all possible choices were explicit, (2) selecting the approach that would force people to make their choices explicit, and (3) selecting the approach that minimizes the number of opt-outs (Thaler & Sunstein, 2003a, p. 178). Such methods are necessary, as the Libertarian Paternalist will have the knowledge and the ability to influence people's decisions. Therefore, for the act to be paternalistic, it needs to be in the best interest of people.

Even though such methods have been developed, still today much discussion about the ethicality of libertarian paternalism and nudging is taking place. At this early stage of the research, we will not go further into discussions about whether or not, or under which conditions, nudging is ethical, but we will save such questions to be discussed at a later stage.

According to Thaler & Sunstein, nudging or paternalistic actions are possible as human beings fail to be rational. As already discussed, most research of psychology and behavioral economy shows how people *"use heuristics that led them to make systematic blunders, exhibit preference reversals and… exhibit dynamic inconsistency [in] valuing present consumption much more than future consumption"* (Thaler & Sunstein, 2003). Such evidence is, as mentioned, a break with the idea of the economic man, Homo economicus, consistent with Simon's research on the bounded rationality, and it legitimizes the thinking of libertarian paternalism. Libertarian paternalists can take advantage of such systematic mental shortcuts that humans make and use them to gently push or '*nudge*' people towards the better choice.

Choice Architecture and Choice Architects

"A choice architect has the responsibility for organizing the context in which people make decisions" (Thaler & Sunstein, 2009, p. 3). Even small and inconsiderable details can have crucial and important impacts on people's behavior. These small details are used to focus the attention of the users in a certain direction (Thaler & Sunstein, 2009, p. 3). The libertarian paternalist is by nature a Choice Architect, and he or she is even a deliberate Choice

Architect who intends to organize the context in which people make decisions in a certain way that influences behavior. When deliberately designing choice architecture, the architect should ensure that his or her actions are paternalistic, in the same sense as the libertarian paternalist seeks to do. This includes e.g. making a cost-benefit analysis or applying one of the three methods for determining whether an approach or a nudge is welfare promoting.

However, not only libertarian paternalists are choice architects. Many people will without even knowing it be choice architects. For example, if you are a doctor and you have to prescribe the best alternative treatments to a patient, you are a choice architect, and if you are a parent who is describing different possible educational options to your child, you are a choice architect (Thaler & Sunstein, 2009, p. 3). Evidently, design is inevitable. All decision-making situations must be designed and structured in some way, no matter whether the setup is deliberately designed to impact decisions in one or the other way.

Nudging by Thaler & Sunstein

In 2008, a couple of years after Thaler and Sunstein first presented the idea of Libertarian Paternalism, they published their book '*Nudge'*, which would show to make nudging a popular concept among the general masses. In this book, they define a nudge as *"any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives. To count as a mere nudge, the intervention must be easy and cheap to avoid. Nudges are not mandates. Putting the fruit at eye level counts as a nudge. Banning junk food does not."* (Thaler & Sunstein, 2009, p. 6). In theory, a nudge should have no impact on the decision-making situation, as no options are removed and no changes to economic circumstances occur. However, in the real world of Homo sapiens (not Homo economicus) the changes do in fact have a predictable effect. This happens as nudging exploits the knowledge of the human errors and heuristics that influence decision-making.

One of the most well known examples of nudging is from Schiphol Airport in Amsterdam. The airport spent a lot of money on cleaning expenses, especially in the men's rooms, as it seemed like the men did not pay much attention when using the urinals, which created a

bit of a mess. The airport then decided to do something about it. They placed a picture of a small black housefly into each urinal, which then created a target for the men to aim at. This small intervention increased the accuracy of the aiming at the urinals and resulted in less of a mess. In fact, by placing this small fly, the airport reduced its spillage by 80% (Thaler & Sunstein, 2009, p. 4). This small intervention and insignificant detail had a major impact on the behavior of the men using the urinals at Schiphol Airport in Amsterdam.

In Denmark, Københavns Kommune have also made use of nudging in order to prevent garbage in the streets of Copenhagen. København Kommune were struggling with too much garbage in the streets, and it cost them a lot of money. Therefore, Københavns Kommune wanted to nudge the pedestrians of Copenhagen into tossing their garbage in the garbage bins. They did this by placing neon green footprints in the direction of the neon green garbage bins (Hansen & Jespersen, 2012). Københavns Kommune made a pilot project to identify whether the neon green intervention did actually have an impact on the behavior of the pedestrians. At first, before putting the green foil onto the garbage bins and placing the green footsteps on the ground, they made an experiment, where they handed out caramels wrapped in paper to the pedestrians in central Copenhagen. After handing out the caramels they had to walk around the area, collecting all the wrappings from the caramels, which were tossed in the streets, in bike baskets or in ashtrays at nearby restaurants. Hereafter, the neon green foil was added to the garbage bins and the footsteps were placed on the ground, and the experiment was repeated (Hansen & Jespersen, 2012). This time, it should be easier for the pedestrians to localize the garbage bins, where they could toss the caramel wrapping paper. The experiment showed that after placing the neon green footprints and the neon green foil on the garbage bins, the discarded caramel paper was reduced with 46 % (Hansen & Jespersen, 2012).

Both these cases are great examples of how interventions in the Choice Architecture can prove to have a great effect on the behavior of people. They show how very simple nudges can be developed from knowledge about the human rationality. Again, it must be emphasized that nudging is not limited to biases and heuristics but as the definition proposes, all aspect of the choice architecture can function as a nudge. However, as this thesis cen-

ters on the availability heuristic and the use hereof in decision-making, theories and literature on nudging by biases and heuristics will have the primary focus.

Nudging health behavior by Vlaev et al.

As mentioned earlier, theories of nudging have been widely developed since Thaler and Sunstein introduced it. Some of those who have most recently added to the general understanding of nudging and how it can be applied is Ivo Vlaev, Dominic King, Paul Dolan and Ara Darzi in '*The Theory and Practice of "Nudging": Changing Health Behaviors*'. Vlaev et al. have developed their theory from many of the preceding notions of nudging from e.g. Hollands et al. and Thaler & Sunstein. However, one of their main arguments is that Thaler and Sunstein's definition of nudging and how to apply it is insufficient (Vlaev et al., 2016, p. 551). Especially when it comes to the topic of health related behavior, on which Vlaev et al. center their research. As mentioned, nudging has often times been used in relation to changing people's consumption to the better, i.e. to the healthier choice. In this area, nudging has proved an efficient tool. This is primarily because traditional tools to improving health and well-being have proved insufficient. The problem is that tools such as policies and interventions build upon the idea that people change their behavior when their motivations and intentions are changed (Vlaev et al., 2016, p. 550), but often times, this is not the case. This is why nudging has proved successful.

According to Vlaev et al., nudging should be used as a tool to more efficiently changing behaviors of health. In this, they consider Holland el al.'s definition to be more thorough. They define nudging as: "*interventions that involve altering the properties or placement of objects or stimuli within micro-environments with the intention of changing health related behavior. Such interventions are implemented within the same micro-environment as that in which the target behavior is performed, typically require minimal conscious engagement, can in principle influence the behavior of many people simultaneously, and are not target or tailored to specific individuals*" (Hollands et al., 2013 in Vlaev et al., 2016, p. 551). Evidently, this definition centers on health-related behavior and adds a spatial element to the definition that dictates how the intervention should be implemented within a specific microenvironment. This builds on the idea that people are greatly influenced by the context and environment in which they make decisions (Vlaev et al., 2016, p. 551). A microenvi-

ronment includes both the physical and social elements of the setting, and the interventions can both include small-scale changes in physical and social elements (Vlaev et al., 2016, p. 551). By having knowledge about both human decision-making and about the decision-making environment, one has the opportunity to influence the choices that people make (Vlaev et al., 2016, p. 551). Hence, according to this definition, a Choice Architect can implement interventions in a certain microenvironment with the purpose of changing health-related behavior.

Nudging by heuristics

From the above introduction to nudging, it is evident that we can nudge people in their decision-making processes by taking advantage of certain heuristics that bias the decision-making process. By being aware of what affects people in the decision-making process, we can nudge them into making *'the better decision'* or *'the rational decision'.* These are decisions that people are often not capable of making themselves because of the complexity of the situation or information overload that the human brain cannot comprehend.

Nudging by the availability heuristic

The availability heuristic is of primary concern of this research, which is why we in this section look into how availability can be used as a means to nudge. According to the research of Kahneman and Tversky, the availability of memories can impact decision-making, as it can lead to inaccurate predictions of probability of an event. The way in which a memory is available to you in a choice situation can be a result of several factors. If trying to impact the availability of certain memories in such a choice situation, one could propose that it is possible to nudge the decision in a certain direction. Looking into the two factors that are of our primarily concern within the availability heuristic; the retrievability of instances bias and the imaginability bias, we propose that such biases can function as a means to nudge decisions. First of all, we propose that if one can make sure that a certain memory is particularly available for a person during the choice situation, one can impact the decision towards a certain outcome. This could for example be by exposing people to a message or an event right before their decision-making. Second, we argue that the imaginability bias can be triggered by making a certain event stand out particularly in memory, which then impacts decision-making. For example, if all positive elements of a certain case are

spelled out or if a message is well presented and well communicated it should make it easier for the receiver to understand and imagine. This will then impact the availability of such event or message in future choice situations.

We have not seen any cases of nudging that have given special attention to this specific bias, which is why this thesis seeks to investigate the potential benefits of applying the availability heuristic as a means to nudge. In the literature, we have only found examples of situations where the availability heuristic negatively impacts decision-making situations, e.g. a case of doctors, who showed to be biased by the availability of recent instances of illnesses when diagnosing patients, which led to systematic failures. This case is presented in the next part. However, as we find that biased decision-making is not always to be overcome, we are of the conviction that the availability heuristic can as well be used in a more positive sense to nudge people to making better decisions.

Nudging habits in System 1

Evidently, nudging theory exploits the cognitive limitations of the human brain and the systematic errors as a result of these. What nudging seeks to do is to improve such systematic decisions that are driven by System 1-thinking and exposed to heuristics and biases. Though, the System 1-thinking has been further developed as researchers have found it to be more complex than thought initially. It has been found that this “System 1” contains at least two different ways of thinking. Therefore, one should consider how nudging works on both of these systems, i.e. both the impulsive and the habitual system. The question becomes whether it is possible to nudge both impulsive actions and habitual actions. When acting out of habit, it is suggested that one might have a tendency to do as one normally does, e.g. in terms of eating lunch at work, one could propose that people do as always no matter the level of hunger or the selection at the buffet. We will give attention to this problem during this thesis, seeking to identify whether the actions that we attend to are habitual or impulsive and what this might mean for the effect of the nudge.

Examples of nudging and biased decision-making

Finally, we present some examples of how nudging, heuristics and Choice Architecture have been used by other researchers to impact decision-making. These examples help us

define the scope of this thesis and they help us in the development of a well-considered and well-crafted experiment. After investigating these experiments, we are able to apply the learning and takeaways from the experiments in the development of our experiment. In this, we seek to ensure that the design is legitimate and able to investigate what we set forth to do. We present two examples, which are found interesting for what we have set forth to investigate. First, we look into an experiment that investigates matters of the availability heuristic and how it can result in faulty decision-making. Hereafter, we present a classic example of a choice architect that designs the architecture of a canteen, which impact the consumption of the users of the canteen.

Here follows the example of how the availability heuristic has been investigated and which effect this has in a real-life experiment. It is the case of the doctors' biased diagnosing, which we mentioned briefly earlier. A number of biases can affect the ways doctors gather and use evidence to provide their patients with the right diagnose. Biases can occur at every phase of the interaction between the doctor and the patient, as the doctor has to estimate the probability of a number of possible diagnoses. This involves deciding which additional information to collect from the patient in order to rule out all diagnoses except one. It ends up with evaluating the gathered information to come up with the correct diagnosis (Bornstein & Emler, 2000, p. 97). When a doctor is in the process of determining a diagnosis, it involves assessing the probabilities of competing diagnoses until the probability of one diagnosis is higher than the others (Bornstein & Emler, 2000, p. 100). The estimation of probabilities can be influenced by the availability heuristic, as the availability heuristic concerns "*estimating of probabilities of an event by how easily one recalls similar events*" (Bornstein & Emler, 2000, p. 100). Recalls should be understood as how '*available*' memories are of similar events. In the case of the doctors, the patient will present a range of symptoms that are typical for one or more diseases. From this, the patient will receive a probable diagnosis. If the patient presents symptoms that are similar to influenza within a period where many people have influenza, the doctor will consider the probability that the patient has influenza very high because of recent events that are easy to recall, and the patient will most likely be diagnosed with influenza (Bornstein & Emler, 2000, p. 101). The doctor could however, mistakenly diagnose the patients because of the availability heuristic, and it could in fact be another and potentially more severe disease. This indicates that

the availability heuristic can lead to a biased probability of assessments, as the availability heuristic "*overestimate probability of a diagnosis when instances are relatively easy to recall*" (Bornstein & Emler, 2000, p 98). In this example, the biased decision-making is clearly faulty and of danger to both the doctor and especially to the patients, which is why such biases should be overcome, as proposed by Kahneman and Tversky.

The final example takes place in a canteen at a school, where the canteen wanted to investigate whether the way the food is displayed and arranged could influence the choices that the kids make without changing the menus (Thaler & Sunstein, 2009, p. 1). They made a range of experiment, where they arranged and displayed the food in different and new ways, and the canteen found that they could nudge the kids towards better and healthier decisions just by changing the Choice Architecture (Thaler & Sunstein, 2009, p. 4). This is a classic nudge in Thaler and Sunstein's thinking. From the experiment the canteen learned that the school kids can be greatly influenced by small changes in the context. The same goes for adults. They will in the same way be very influenced by the Choice Architecture, as proposed by Thaler and Sunstein (Thaler & Sunstein, 2009, p. 2).

With these examples we finalize the review of previous literature and research within the field of behavioral economics and decision-making. From this we can move forward with our own research. Therefore, we continue with the hypothesis development in the next chapter, which has the purpose of guiding us in our experiment and the research in general.

HYPOTHESIS DEVELOPMENT

The focus of this chapter is to develop relevant hypotheses based on the literature review that can guide the research and help answer the research question. We seek to develop several hypotheses that can all be tested in the experiment that we conduct. The experiment will take place in a canteen over a period of four days and it includes nudging two separate experimental groups at two different points of time. In the experiment, we present the two groups of people with a single survey question prior to their lunch, asking about their consumption habits within the category of fruit and vegetables. The following decision-making that we seek to impact takes place in the canteen at the lunch buffet. After the lunch, the respondents are met with another question that asks about the actual consumption. The experiment is presented in further detail in the methodology chapter.

In the following, seven hypotheses are formulated. They are based on the theoretical findings that we presented in the literature review, and in this sense, they are an outcome of fundamental theories within the field of behavioral economics and marketing as well as of previous research within these fields. With the hypotheses, we seek to further develop the understanding and exploitation of nudging by heuristics and biases from a marketing and communication perspective.

The research question of the thesis is the point of departure. The research question seeks to investigate: *"How can the availability heuristic be used as a nudge to impact decision-making, and how can one isolate and measure the effects of a specific heuristic? Further, what are the consequences of the uncritical and general use of nudging in mass society?".* From this we formulate the overall hypothesis:

> *H1: Participants stating their expectations for consumption behavior on consumption of fruit and vegetables will result in at least the stated consumption or an increased consumption of fruit and vegetables.*

The overall hypothesis is based on the research of heuristics from Kahneman and Tversky and Vlaev and Dolan. First of all, Kahneman and Tversky propose that people often assess the probability of an event to occur based on the ease with which previous instances

come to mind. Further, the ease with which one can imagine the event also impacts the judgment of probability (Tversky & Kahneman, 1974, p. 1127). Second, stating a desired goal either verbally or written will make people feel more committed to this goal (Vlaev & Dolan, 2015, p. 85). As previous research has shown that it is possible to change people's behavior by making small changes in the choice situation, we propose that by adding a small social element to the choice architecture, we can impact the decision-making process. Especially within the area of health, nudging has proven very useful in changing people's consumption habits to the better. Therefore, we propose that asking people to state their expectations for consumption will first of all commit them to a certain goal, and second, it will ensure that stated goal is available to the participants during the choice situation. Evidently, the first hypothesis contains elements of both the availability heuristic and the commitment heuristic.

As the thesis seeks to investigate the effects of the availability heuristic and hence sets forth to isolate the availability heuristic from other potential heuristics, we develop a set of sub-hypotheses that help us investigate the different heuristics across a three variables. These variables are days, genders and groups. The variables and the experimental design is presented in detail in the methodology chapter.

With the three variables we seek to investigate how and if groups (time of day), genders and days impact the results of the experiment, i.e. if groups, days and genders impact consumption. In this, we work with two observations: expected consumption (EC), which corresponds to the statements made by the participants prior to their lunch, and their statement of actual consumption (AC) after their lunch. Expected consumption is the independent variable and actual consumption is the dependent variable (see figure 2). During the experiment, we seek to test how time, day and gender moderate the effects of the nudge, which is illustrated by the arrow from expected consumption to actual consumption. Evidently, groups, days and genders are moderators. A moderator is a variable that changes the relationship between the dependent and independent variable.

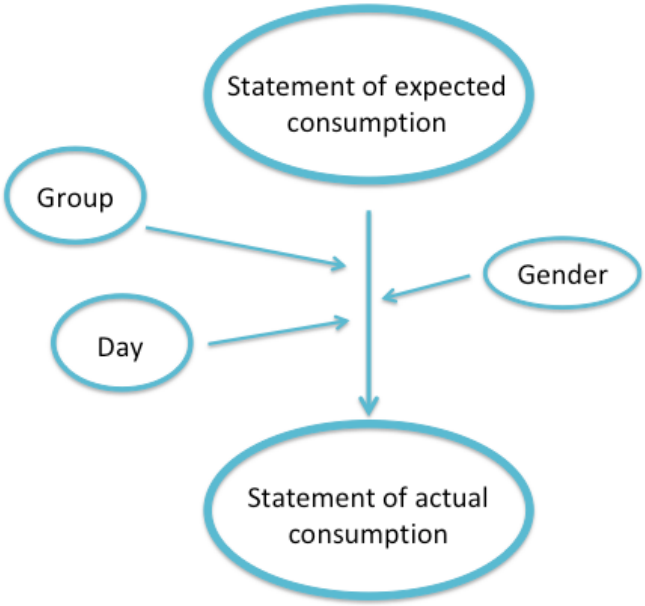


Figure 2: Experimental variables 1.0. Source: Authors' own work.

This leads us to develop a set of hypotheses that suggest how the three different variables might impact the results of the experiment. Therefore, six sub-hypotheses are developed to broaden the understanding of the problem. Following the logic of the overall hypothesis, the first sub-hypothesis is:

> *H2: The nudge will have a larger effect on group 2, who has most recently stated the expectation for consumption, than on group 1, i.e. there will be a larger increase in consumption for group 2 than for group 1.*

Based on the understanding that recent occurrences will often be more available and hence have a larger impact on the judgment of probability than earlier occurrences, this second hypothesis is formed. From this hypothesis, we can come closer to an understanding of how the retrievability bias in isolation has an impact on the decision-making process, and hence how group as a variable moderates the result.

Next off, we seek to understand if there are any differences in the data caused by genders. This leads to the second sub-hypothesis, which builds on the element of social desirability. In line with Thaler and Sunstein's notions on optimism in social contexts, it is believed that the nudge will have a larger impact on certain groups of people, as people often

behave differently in social contexts. This could be ascribed to the social desirability bias, which is the tendency for people to present themselves in the best possible light. It is proposed that certain people, who to a larger extent are biased by social desirability, might be easier to nudge (Fischer, 1993, p. 303). Social desirability is especially important to take into consideration when treating a sensitive topic. It is argued that healthy eating can be a sensitive topic for certain people and it is as well a topic that is widely discussed in society today. It is well-known that eating fruit and vegetables is good for your health and there exist clear values for what is healthy and what is not. Therefore, taking the topic of this experiment into mind, it is suggested that women might be more affected by the social desirability bias than men, based on an assumption that women care more about healthy eating than men.

> *H3: Women will be more affected by the nudge than men, i.e. the increase in consumption will be greater for women than for men*.

In this thinking, we suppose that the participants believe that a socially desirable answer is a healthy answer and that this will result in them being seen more favorably by others. Therefore, being affected by the nudge is understood as an increase in the consumption of fruit and vegetables.

The hypotheses 4 and 5 are concerned with the commitment heuristic and the degree to which committing oneself to a goal impacts the decision-making. We propose that the power of the commitment heuristic varies across groups, i.e. the different times of the day, and across the four different days. In this, it is proposed that days and groups moderate the effect of the nudge.

> *H4: People are more inclined to keep a recent promise of consumption (group 2) than an earlier promise (group 1), i.e. the difference between EC and AC will be larger for group 1.*

> *H5: People are more inclined to keep their promise of consumption in the beginning of the week, i.e. EC is closer to AC in the beginning of the week*.

The next two hypotheses are based on theories about the optimism bias and thereby overconfidence from the participants. The optimism bias is defined as the difference between one's expectation and the actual outcome that follows (Thaler & Sunstein, 2009, p. 32), in this case being the difference between expected consumption and actual consumption. Again, the hypotheses propose that there are differences in optimism across the two groups, i.e. the different times of the day, and the different days:

> *H6: People overestimate healthy consumption in the beginning of the day, i.e. the difference between EC and AC will be larger for group 1 than for group 2.*

> *H7: People overestimate healthy consumption in the beginning of the week, i.e. the difference between EC and AC will be larger in the beginning of the week.*

With the hypotheses 6 and 7, we suggest that the variables 'day' and "group" might in fact also function as independent variables that impacts the level of expected consumption directly. If this is so, the statement of expected consumption becomes a dependent variable, and the statement of actual consumption is the second dependent variable. This suggests that the two variables can both function as moderators and as independent variables, as both group and day can impact first the statement of expected consumption and at the same time moderate the relationship between EC and AC (see figure 3).

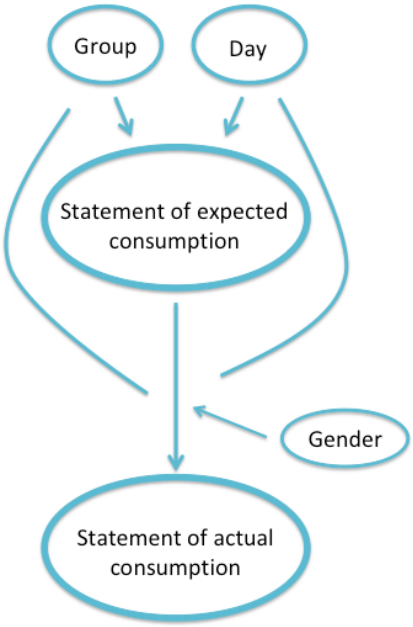


Figure 3: Experimental variables 2.0. Source: Authors' own work

From the seven hypotheses, it is evident that we as researchers function as Choice Architects in the experiment. With a small intervention in the choice architecture, we seek to impact the behavior of the participants by the use of a nudge that builds upon a set of heuristics and biases. As the nudging theory defines, we as Libertarian Paternalists should seek to facilitate the rational decision with the nudge to ensure that we make the participants 'better off'. Evidently, what is in general considered to be 'better off' in the seven hypotheses is an increase in consumption of fruit and vegetables. It cannot be known whether eating a larger amount of fruit and vegetables is in fact the rational choice that will make all the participants 'better off', or whether in some instances, eating exactly the amount of fruit and vegetables that one sets forth to do or even eating less, is in fact a healthier choice in terms of e.g. nutrition and well-being. As presented in the literature review, Thaler and Sunstein have developed a set of methods for how to come closer to a truthful determination of what 'better off' actually is. We attend to this in the development of the experiment. Further, the general issue of what 'better off' means is again brought up in the discussion.

The seven hypotheses are tested in the analysis of the data that we collect through the experiment. From this we either accept or reject the hypotheses based on the findings of the analysis, which leads to a discussion of the results with the purpose of broadening the understanding of how the availability heuristic can function as a means to nudge.

METHODOLOGY

This chapter introduces the methodological approach taken to the thesis. It starts with considerations about philosophy of science, followed by a presentation of the research design, which includes a thorough review of the experimental design. Finally, it concludes with an overview of the data measurement.

Philosophy of Science

The purpose of this section is to determine the position of this research in terms of philosophy of science, i.e. what scientific approach, the paradigm that is taken to the research question. This is important because the scientific approach, the paradigm, defines the appertaining ontology and epistemology, which are determinant for the methodology.

A paradigm is defined as *"a set of fundamental principles that guide a field of research"* (Presskorn-Thygesen in Nygaard, 2013, p. 23). Philosophy of science is concerned with two primary paradigms, the positivistic and the constructivistic. The positivistic and constructivistic convictions or paradigms are extremes and in between these two most research is found. The same goes for this thesis and we propose that the thesis belongs to the social constructivism paradigm. To a paradigm belongs a certain ontology and epistemology. Ontologi is *'the study of the nature of being'* (Nygaard, 2013, p. 13) and epistemology is *'the theory of knowledge'* (Nygaard, 2013, p. 13). The ontology can be either realistic or relativistic (or something in between), and the epistemology is either subjective or objective (Presskorn-Thygesen in Nygaard, 2013, p. 28).

Within positivism the ontology is realistic where the researcher sees the world as something that exist independently of us. The epistemology is objective and within positivism one believes that an objective truth can be found from observations of the world. The methodology is quantitative with the purpose of measuring and observing the world within which we find ourselves (Presskorn-Thygesen in Nygaard, 2013, p. 29). Constructivism is the complete opposite. Within this thinking, the ontology is relativistic, i.e. the researcher believes that the world is a construction of social constructions. The epistemology is subjective as knowledge is bound by social and subjective constructions. Constructivists believe that an objective truth is impossible. The methodology is qualitative with the purpose

of uncovering the problems that society considers as objective truths (Presskorn-Thygesen in Nygaard, 2013, p. 29).

We argue that this research belongs to the constructivistic paradigm. What is interesting about constructivism is that the ambition of the paradigm is 'double'. Presskorn-Thygesen claims that constructivism on the one hand seeks to show that common understandings that seem to be 'true', 'natural' and 'inevitable' are not objective and that these are in fact social constructions. On the other hand, constructivism seeks to map out the mechanisms of the social construction processes that result in that some common understandings are shared in a social community and come to seem, as if they are objective and true (Presskorn-Thygesen in Nygaard, 2013, p. 37).

In general, the approach to the areas of psychology and marketing is rather constructivistic. Actually, it is often considered to be social constructivistic representing the belief that reality is a social construction dependent on the participating parties. Based on this and on the nature of the research question of this thesis, we find that this research belongs to the social constructivist paradigm. According to Presskorn-Thygesen, the constructivistic paradigm has a *relativistic ontology* where the scientific objects are a direct consequence of the research observation and the reality is primarily made up of social constructions. The *epistemology* is *subjective* as no objective truth is possible when reality is a subjective construction (Presskorn-Thygesen in Nygaard, 2013, p. 29). From this research philosophy, taking a *qualitative methodology* to answering the research question will be most appropriate.

In this thesis, we seek to understand the psychological and social mechanisms behind biases, heuristics and nudges that make reality, or in this case decision-making, seem objective. Nudging is per definition an idea that springs from the social constructivist thinking, where one believes that small interventions construct the reality that we experience as 'true'. In an objective world or in a positivistic world, such nudging-interventions should not have an effect but in the social reality they do. This is again the thinking of Homo sapiens instead of Homo economicus in decision-making theory. Therefore, we as researchers adopt the idea of social constructivism believing that the availability heuristic will affect reality, as it is experienced by the participants in the experiment that we set forth to conduct.

Evidently, we observe the reality of the participants from a relativistic perspective. It is expected that a statistical coherence can be found in the data collected through the experiment, i.e. that this impact that we expect to observe can be measured. This implies that there is in fact order in the social reality and that a common understanding is shared as being true. Therefore, it is believed that there is not one true reality and not one objective truth, but rather that many subjective realities can exist and that realities are created from current circumstances. However, it is expected that coherences in the reality can be found in objective data, and from this we will be able to deduce some general truths. This proposition is in line with the double ambition of the social constructivist thinking.

Research Design

We take a deductive approach to the research where we start out by looking into previous theories and literature that lay the foundation for this research. From this, we deduce a set of hypotheses that guide the research. We seek to accept or reject the hypotheses based on the primary data that we collect through the experiment. All the data collected in the experiment is analyzed and from this data we can test the hypotheses. Finally, we discuss the findings with the purpose of answering the research question that we set forth to investigate.

The working approach is qualitative and hermeneutic, though seeking to determine some seemingly objective truths about reality. As we move forward in the research, our cognition and realization is developed. We move from a pre-understanding to an understanding in a continuous loop where pre-understanding sparks discussions that move us to a new understanding. It is acknowledged that we as researchers have a pre-understanding that inevitably impacts the research. However, by making ourselves aware of our pre-understanding we seek to minimize how such prejudices and pre-understandings might negatively impact the research.

Experimental Design

In the following, we describe the experimental design in depth. The section includes a discussion of multiple considerations and possible implications of the choice of experimental design.

Developing the experimental design

The experiment is constructed on the basis of the hypothesis development, where a set of hypotheses is deduced from the theories and previous research within the field of study.

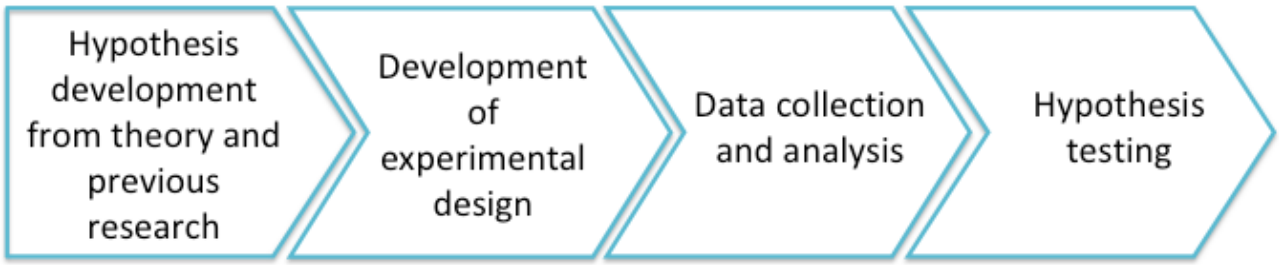


Figure 4: Experimental design. Source: Authors' own work.

The purpose of the experiment is to test whether one can use the availability heuristic to nudge people's behavior. It seeks to understand how an easily recollected piece of information can impact decision-making. The experiment is conducted in a canteen, and the participants of the experiment are users of the canteen. We as researchers take on the role of Choice Architects in seeking to impact people's behavior by the means of a nudge.

Validity and reliability

It is essential to briefly comment on the validity and reliability of the data collected through the experiment to verify whether the findings are valid and trustworthy. Only if the applied method is valid and if the data can be reproduced, hence if it is reliable, will we be able to test the hypotheses and make valid conclusions on the matter. Kvale defines validity as *"the degree that a method investigates what it is intended to investigate, to "the extent to which our observations indeed reflect the phenomena or variables of interest to us""* (Kvale, 2013, p. 122). Reliability then refers to *"the consistency and trustworthiness of research findings: it is often treated in relation to whether a finding is reproducible at another time and by other researchers"* (Kvale, 2013, p. 122).

To collect valid data and conduct the experiment as objectively and reliably as possible, it is important to reduce the number of biases that can have an impact on people's consumption. This is to ensure that the data that we collect is useful for the problem area that we investigate. We have defined a range of biases and heuristics that might impact the results but we are aware that this list might not be the full list of potential biases that impact the decision-making. Further, we acknowledge that our mere presence in the canteen impacts the behavior of the participants but we seek to keep this at a minimum. For example, by staying right outside the canteen, we seek to minimize the impact as the participants are not constantly reminded of our presence. By doing so, we seek to ensure that the experiment only investigates what it is intended to do. In this case: the effects of the nudge.

An important point to make when it comes to the presence of potential factors that impact the results is the fact that the price is not a moderator. The users of the canteen pay a monthly fee to eat lunch in the canteen and can eat as much as they like and on as many days as they like. Therefore, we find that the only moderators are the ones that we have already defined; groups, days, and genders.

As we outline the experimental setup in detail below, we are of the conviction that other researchers can easily replicate the experiment and supposedly to a great extent get the same result. Hence, it is argued that the research is reliable.

Experimental setup

When developing the experimental setup, we use the definition of nudging by Hollands et al. as our starting point: "*intentions that involve altering the properties or placement of objects or stimuli within micro-environments with the intention of changing health related behavior. Such interventions are implemented within the same micro-environment as that in which the target behavior is performed, typically require minimal conscious engagement, can in principle influence the behavior of many people simultaneously, and are not target or tailored to specific individuals*" (Hollands et al., 2013 in Vlaev et al., 2016, p. 551). We seek to add an intervention to a certain microenvironment to impact people's decision-making. All these elements are presented in this section.

The setup of the experiment is comparative. In a comparative experiment one will have two or more groups who are treated differently. According to Easterling: "*The purpose of doing the experiments is to compare two or more ways of doing something… In this suite of experiments, different experimental units are subjected to different treatments. Responses of the experimental units to the different treatments are measured and compared to assess the extent to which different treatments lead to different responses*" (Easterling, 2015, p. 11). The building blocks of a comparative experiment are (1) Experimental units, (2) Treatments, and (3) Responses (Easterling, 2015, p. 11). In the experiment, treatments are applied to different experimental units and responses are collected and measured. In the following, the three building blocks of the experiment of this thesis are defined.

Experimental Units

The experimental units are "*the entities that receive an independent application of one of the experiment's treatments*" (Easterling, 2015, p. 13). In this experiment, we have two experimental units consisting of randomly selected participants, who are all guests of the canteen. The participants of the two groups are homogeneous but they are treated differently, i.e. that differences in responses should be caused by treatments and not based on difference among the participants. Group 1 is the '*9 o'clock group*' and group 2 is the '*12 o'clock group*'.

Treatments

The treatments are "*the set of conditions under study*" (Easterling, 2015, p. 13). In this experiment, the treatment is a nudge and the nudge takes the form of the question about expectations for consumption posed at different points in time. The two different times are around 9 o'clock as the participants arrive for work and around 12 o'clock as they go to lunch.

The 9 o'clock group

The participants of the 9 o'clock group are presented with two questions and a set of pictures showing five different plates (Appendix 1). The different plates show either a full plate of fruit and vegetables, a plate with ¾ of fruit and vegetables, ½ a plate, a ¼ plate and an empty plate. First, the participants are asked: "*Are you going to eat lunch in the*

canteen today?" *(*in Danish: *"Spiser du frokost i kantinen i dag?")*. If the participants are not going to eat lunch in the canteen, they are not included in the experiment. Otherwise, if the participants answer yes, we ask them the second question: "*Based on these five plates, which one do you consider best shows your consumption of fruit and vegetables at lunch today?*" *(in Danish: "Baseret på disse fem tallerkener, hvilken mener du så bedst beskriver dit forbrug af frugt og grønt til frokost i dag?"*). From this, the participants have five possible choices (Responses), i.e. 0, ¼, ½, ¾, and a full plate of fruit and vegetables. We note their answers on a printed excel chart (Appendix 2) to keep track of the different answers. The participants are given a piece of paper with a number on in order for us to track the data of the participants before and after their lunch. Hereafter, they are told to come back to us after their lunch for a follow-up question. After lunch, the participants are asked about their actual consumption of fruit and vegetables. Specifically, we ask the question: "*Based on these five plates, how will you characterize your actual consumption of fruit and vegetables during this lunch?*" *("Baseret på disse fem tallerkener, hvilken beskriver så bedst dit reelle forbrug af frugt og grønt til frokost").* Again, their answers are noted on the chart. We believe that there will be a minimum dropout, as the participants have to pass us on their way out of the canteen, but we realize that dropouts can occur, e.g. if the participants decide not to eat in the canteen anyway or if they forget about their participation.

The 12 o'clock group

For the 12 o'clock group, we follow the same method as for the 9 o'clock group. When the participants go to lunch, we ask them two questions and show the same five pictures of the plates as the 9 o'clock group. The participants are asked right before entering the canteen: "*Are you on your way to lunch?*" *(*in Danish: *"Er du på vej til frokost?")*. If the participants answer no to the first question they are not included in the experiment. Otherwise, if the participants answer yes, we ask them the second question: "*Based on these five plates, which one do you consider best shows your consumption of fruit and vegetables at lunch today?*" *(in Danish: "Baseret på disse fem tallerkener, hvilken mener du så bedst beskriver dit forbrug af frugt og grønt til frokost i dag?"*). Hereafter, we proceed with the same steps as with the 9 o'clock group. Also for this group, we find it necessary to let the participants know that we have a follow-up question after lunch to try to keep the dropout at a minimum, even though this might impact their answer to the questions.

With this treatment, we investigate the effects of the availability heuristic and test whether this impacts decision-making. As the participants make their statement about expected consumption, they verbally commit themselves to an action. Further, the use of the five pictures with fruit and vegetables should make the action easier to image. We propose that this strengthens the effects of the nudge. Hence, this treatment functions as the small-scale intervention that should impact the behavior of the participants. We argue that the intervention alters or adds a stimulus to the microenvironment, in this case the canteen, with the intention of changing behavior.

Responses

The responses are "*the measured characteristics used to evaluate the effect of treatments on experimental units*" (Easterling, 2015, p. 13). In this experiment, the responses are the observations of amounts of expected and actual consumption on the 5-step scale that goes from 0 to 1, including the observations 0, ¼, ½, ¾ and 1.

To sum up, we seek to investigate how asking people to state their expectations for consumption can impact their actual consumption, i.e. how stating an expectation or a conviction can function as a recent piece of information that is easily retrieved in the decision-making situation (retrievability bias) and hence impacts the decision. In order to understand the power of the availability heuristic and whether it impacts decision-making, we collect responses from the two different experimental units, who are treated with the same nudge but at two different points in time. Group 1 is nudged around 9 o'clock and group 2 is nudged right before going to lunch around 12 o'clock. The purpose of doing so is to determine whether the different times of treatment and hence the availability heuristic has an effect, or whether the potential changes in behavior must be ascribed to other biases as presented in the hypothesis development. We propose that the nudge exploits both the availability heuristic, including imaginability and retrievability biases, as well as the commitment heuristics. Further, we acknowledge that the social desirability bias and the optimism bias may further have an effect, as proposed in the hypothesis development. With the experiment and the following data analysis, we seek to analyze the data across days, groups, and genders to determine exactly which biases come into play at which point in time and why, with the purpose of isolating and measuring the effects of primarily the availability heuristic.

Data collection

In this section, we describe the methodological considerations and observations before and during the data collection. First, we present a review of the preliminary considerations prior to the data collection. These are followed by a description of the population, before moving on to the actual data collection. Finally, we list possible limitations of the experimental design.

Preliminary considerations

The experiment takes place during four days in February from Monday till Thursday in the canteen in the ISS Tower in Søborg, Denmark. Prior to conducting the experiment, we contacted the canteen manager and got the acceptance to carry out the experiment and he agreed to help us with regards to the selection of food in the canteen. As mentioned, the canteen is the microenvironment within which the experiment takes place and which we seek to alter with a small-scale intervention. The canteen and its buffet are the elements that are used to illustrate the effect of the nudge in the experiment. The food selection in the canteen includes a daily salad buffet, a selection of cold meats, a warm dish of the day and occasionally soup. Each Wednesday, the dish of the day is fish. The rest of the week it can be chicken, pork, beef and lamb. There is typically also a vegetarian or a vegan dish of some sort. The selection of cold meats varies from day to day. On Thursdays, there is a dessert/cake. To overcome changes in consumption caused by the difference in food selection, the experiment is carried out over a period of four days. During these four days of the experiment, the canteen agreed to avoid adding too much meat, beans, and especially pasta and rice on the salad buffet in order for the participants to better be able to categorize the actual consumed amounts of fruit and vegetables. We aim to collect the same amount of data across each of the four days that the experiment runs. This is to ensure that the data is not reliant on any specific day and hence food selection. We aim to collect data from approximately 20 participants from group 1, 30 participants from group 2 and approximately 15 from the control group each day. The participants are chosen randomly.

Prior to conducting the experiment, pictures of five plates are taken to show the different amounts of fruit and vegetables. As mentioned, the different plates show either a full plate of fruit and vegetables, a plate with ¾ of fruit and vegetables, ½ a plate, a ¼ plate and an

empty plate. We choose to visualize the consumption in this division of the plate to make it easier for the participants to estimate their amount of fruit and vegetables consumed. The visualization should make the decision situation more accessible and imaginable for the participants, as they will be able to see exactly what the different amounts look like. The pictures of the plates are taken in the canteen where the experiment is conducted and they show vegetables and fruit from the actual buffet to keep the identification and recognition at a maximum and to make the decision-making situation as realistic as possible.

We choose to divide the plate into four parts, which is to some extent inspired by the Danish Veterinary and Food Administration's (part of the Ministry of Environment and Food) recommendation for food division. This division is called the Y-plate (Appendix 3) and it is used as a recommendation for one's daily consumption. The Y-plate divides the plate into three parts; two times 2/5 and one time 1/5 of the plate, where fruit and vegetables as well as carbs should each make up 2/5 of the plate and protein should make up the last 1/5 of the plate. On this basis, it is expected that the participants can easily determine the expected and actual consumption based on the pictures, as this is common practice when discussing consumption. A similar guideline from The Veterinary and Food Administration recommends eating 600 grams of fruit and vegetables each day but asking people to estimate the amount of grams that they expect to consume at lunch would be too complex a task for most people to perform (Fødevarestyrelsen, 2018). However, selecting visually how much fruit and vegetables they expect to consume is more intuitive and simple, which is why we chose to include the pictures in the experiment.

When asking the participants to estimate their consumption of fruit and vegetables, they will not have access to all the necessary information to make a fully rational decision. For example, most of them might not even know what they will find on the salad buffet or what the dish of the day is. Therefore, we argue that they will make a bounded rational decision based on the information that is available to them at this point in time. Therefore, we believe to get answers like "*I think I will eat…*" or "*It is likely that I will eat….*" from the participants. In this, we suppose that many of the participants probably have some sort of idea of how much vegetables and fruit they will consume. The participants will base their decision on a subjective probability, as it is likely that the participants will make their decision and act based on previous subjective events in the canteen.

As a final element of the preliminary considerations, we shortly discuss whether the act of enforcing the participants to eat a larger amount of fruit and vegetables can actually be characterized as a paternalistic act as proposed by Thaler and Sunstein. To determine whether eating more fruit and vegetables do in fact make the participants 'better off', we attend to one of the three methods that Thaler and Sunstein propose can assess whether the act is found to be paternalistic. If applying the method of "*selecting the approach that the majority would select if all possible choices were explicit*", we propose that the majority would eat at least the amount of fruit and vegetables as the Danish Veterinary and Food Administration recommends, i.e. 2/5 plate of fruit and vegetables. Further, as the majority is of the conviction that fruit and vegetables are healthy, we propose that the majority would in general eat a considerable amount of fruit and vegetables if all possible choices were explicit. Therefore, we consider our act, i.e. the intervention made in the experiment, as being paternalistic, as it sets forth to increase the level of fruit and vegetables consumed by the participants. After conducting the experiment, we look further into the paternalistic element of the intervention.

Population and randomization

In the following, we present a demographic description of the population followed by considerations about the randomization of the participants of the experiment.

The population consists of all the tenants and their employees at the ISS Tower, who eat in the canteen on a daily basis. The population consists of 554 people with an age range of 18 to 67 years old. There are 348 men and 206 women, with 13 different nationalities, and their occupation ranges from Student Assistants to CEOs. The tenants are from 12 different companies that operate within different industries. In the experiment, we seek to include a wide range of the workforce. The participants of the experiment can count both student assistants, chief executive, technicians, cleaning staff, receptionists and different administrative employees. With a well-varied group of participants, it is argued that the results are representative for the whole workplace and not limited to a certain type of employee or organization.

With a population of 554 employees, we calculate the sample size based on a confidence level and margin of error. Having a confidence level of 95% and allowing for a margin of

error of 6% we will need a sample size of 180 participants in the experimental groups (Creative Research Systems, 2012).

The time that the employees arrive for work in the morning and the time when they go to the canteen for lunch is random. The canteen is open from 11.00 till 13.30. Therefore, including participants randomly in the order they arrive in the morning, or picking out participants as they go to the canteen for lunch, counts as random selection. When using random selection of participants in both experimental groups and in the control group, there are risks of having more male than female. However, this is taken into account during the data analysis and it is not expected to have a critical impact on the results.

Collection of data from the experimental groups

We collect data from the two experimental groups across all four days. For the 9 o'clock group we collect data on expected consumption in the morning as they arrive to the building and we collect data on the actual consumption after their lunch break. For the 12 o'clock group we collect data about their expectations for consumption right before they go to lunch, and after their lunch break we collect data on their actual consumption. Additionally, we include a control group from which we collect data on actual consumption after the lunch break. The participants of the control group are shown the same five pictures of the plates and they are asked the same question as the participants of the experimental groups after they have eaten their lunch. The participants in all three groups are randomly selected among the users of the shared canteen at ISS Tower. For group 1, we collect data from approximately 80 participants *(note: Monday, (day 1) we asked 20 people, but as we experienced a relatively large drop-out (25%) in group 1 we decided to increase the number of participants per day to 25 participants)*. For group 2, we collect data from approximately 120 participants, and the control group will consist of 60 participants. We are aware that possible dropouts occur.

With regards to the validity of the data, we realize that it would be ideal if the participants did not know that they were going to be asked about their actual consumption afterwards. As mentioned, we find it necessary to let the participants know that we have a follow-up question for them after the lunch. This is to keep the dropout at a minimum. If the participants did not know that we want to ask them another question after their lunch, they would

believe that their participation is limited to the statement about consumption before going into the canteen, and as they do not expect to be asked about actual consumption afterwards, we believe this would result in the minimum impact on actual behavior. Potential changes in consumption can then be fully ascribed to the different heuristics. However, this will not be possible in practice. Due to the heavy flow of people in the canteen, we will not be able to keep track of all the participants included in the experiment, and there will be larger risks of dropouts. Therefore, it is necessary to let people know that we have a follow-up question. However, to keep this at a minimum, we give the participants a piece of paper with a number on and tell them to come back to us after lunch, without letting them know what we are going to ask them. Then, as the participants exit the cantina, we ask them about their actual consumption.

As presented earlier, we apply the same treatment to both experimental units, i.e. the same intervention for both groups is implemented in the microenvironment of the canteen. However, one could claim that the two experimental groups are in fact found in two separate microenvironments. Even though the spatial environment is the same for both groups, there is a temporal element, which should be taken into consideration. As we ask one group around 9 o'clock as they arrive for work and a second group right before lunch, one could assume that the participants of the two groups find themselves in two different temporal microenvironments and therefore have two different mindsets even though the physical environment is the same. This is taken into consideration in the analysis of the data.

Limitations of the experiment

Within all three groups, the 9 o'clock group, the 12 o'clock group and the control group, there is a slight limitation in the collection of data for the actual consumption. We cannot know if what the participants say is true but we will have to believe that what they say they ate is in fact what they actually did eat, even though it includes making an estimate that might vary from participant to participant. What participant X considers a full plate of vegetables, might not be the same as what participant Y considers a full plate. Some might find that chickpeas, dressings and other foods found on the salad buffet is included in the amount of vegetables, whereas others might exclude this. Though, as mentioned, the canteen manager agreed to keep the amount of beans, meat, and pasta on the salad buffet at

a minimum during the days of the experiment. Further, as an attempt to overcome this limitation we make use of the five pictures, which make it easier for the participants to actually select the exact amount that best matches their actual consumption. Therefore, we must believe that what the participants tell us is true and that the picture they choose is almost an exact picture of what they actually consumed.

It could be argued that the participants of the experiment may be giving answers that they think are considered 'the correct answer', e.g. the healthier answer. One could then expect that the participants will state that their consumption is higher than it actually is, in order to seem healthier or better. This is based on the assumption that everyone knows that eating fruit and vegetables is good for your health, and also on the commonly known guidelines from the Danish Veterinary and Food Administration. One way to test for the truthfulness of the participants' answers is to test for the level of *'social desirability'* among the participants. According to Crowne and Marlowe, it is possible to measure the social desirability of people, i.e. their need to present themselves in a socially desirable/undesirable way. The level of social desirability can be determined on a scale, based on a set of questions concerning personal attitudes and traits. Knowing of a participant's position on this scale will enable an assessment of whether or not the results of the experiment are impacted by a high degree of social desirability, and from this, one will further be able to assess the reliability of the results (Crowne & Marlowe, 1960, p. 350).

Marlowe and Crowne's Social Desirability Scale lists 33 statements and from answers of true and false one will be able to determine the level of social desirability of the participants. Amongst these statements are e.g. (6) *"I sometimes feel resentful when I don't get my way"*, (17) *"I always try to practice what I preach"*, and (31) "*I have never felt that I was punished without cause"* (Crowne & Marlowe, 1960, p. 351).

We choose not to check for social desirability of the participants for a number of reasons. First of all, we find it most important to include as many participants as possible in the experiment, and expanding the experiment to include questions about personal attitudes and traits would naturally have a negative impact on the number of participants that we can manage to include. Furthermore, because we question people during their lunch break on a supposedly busy workday, we seek to keep the involvement of the single participants at

a minimum, making sure not to bother them too much in their daily routines. Including questions of this type could make people reluctant to participate in the experiment. Finally, the 33 statements are of personal character and would be slightly inappropriate to ask people without a warning. However, by being aware of the existence of the social desirability scale and how it might have an impact on some of the results of the experiment, we can take it into account when analyzing the data. Also, we include social desirability in the set of hypotheses that we seek to test and in this seek to determine whether or not it proves to have a critical impact on the results of the experiment.

A final limitation of the experiment is that we cannot know whether we include only the healthiest people at the ISS Tower in the experiment, or the other way around, whether we ask only the less healthy people in either of the groups. However, since we set forth to ask at least 180 people working at the ISS Tower, which corresponds to 32% of the entire workforce at the ISS Tower, we argue that we overcome this uncertainty and succeeds to include both very healthy and less healthy participants. In other words, as we ask such a large percentage of the users of the canteen at the ISS Tower, we believe that the data is representative for the whole workforce.

Measurements

This final part presents the measurements that lay the foundation for the data analysis. The program SPSS is used for statistical analysis of the data of expected and actual consumption that we collected through the experiment. The analysis includes descriptive analyses and comparisons of means with the use of independent t-tests that compare the data across the different variables. We observe the consumption across groups, days, and genders and we look for immediate differences as well as statistically significant differences across all the variables. This enables us to determine if and how the nudge succeeds to impact the behavior of the participants.

Data transcription and data coding

First, the collected data is transcribed digitally into excel in the format of X's marking first gender, then expected and actual consumption on the scale of 0-1 and divided across groups and days. Afterwards we transform the data into measurable values that can be

analyzed statistically in SPSS (Appendix 4). To be able to statistically analyze the observations, the data is coded into values that are usable in the SSPS program. The experiment gives the following five variables: gender, expected consumption (EC), actual consumption (AC), groups, and days. The coded observations are found in the table below:

Men	1
Women	2
EC = 0	1
EC = 1/4	2
EC = 1/2	3
EC = 3/4	4
EC = 1/1	5
AC = 0	1
AC = 1/4	2
AC = 1/2	3
AC = 3/4	4
AC = 1/1	5
9 o'clock group	1
12 o'clock group	2
Control group	3
Monday (day 1)	1
Tuesday (day 2)	2
Wednesday (day 3)	3
Thursday (day 4)	4

Table 1: Coding of observations.

All observations are accumulated and inserted into the SPSS program.

As mentioned, we use the SPSS program for the statistical analysis of the collected data. The SPSS program creates different statistics for the set of variables from which we are able to analyze and conclude upon the data. The statistics that we find relevant to consider in the analysis of the collected data are: *descriptives* with a focus on the different calculated means and standard deviation in order to compare the means, and the *independent t-test* with the focus on the Sig.-/p-value. The following statistics are explained from the definitions given by the SPSS program:

The *descriptive test* determines summary statistics for several variables. The statistical variables, which are relevant for our analysis, are the different means and the standard deviation.

The *mean* is a measure of the central tendency and the mean procedure calculates the mean of one or more independent variables.

The *standard deviation* is a measure that is used to quantify the amount of variation in the data set or the dispersion of a set of data relative to the mean. If the data points from the data set are far away from the mean, there is a higher deviation among the data set.

The *independent t-test* tests the significance of the difference between two sample means. The independent t-test determines whether there is a statistical significance between the means of two independent sets of data, e.g. if there is a statistically significant difference between the grades of one defined group versus another defined group, or between one point in time versus another point in time.

The *significance level (Sig.)* and the *p-value* are two sides of the same value (1-Sig. = p). It expresses the significance of the t-test, i.e. whether or not the immediate differences between two means are statistically significant and not due to chance. For a difference in means to be statistically significant, one should be able to say with 95% confidence that the observable differences are in fact due to a significant difference between the two data sets. Therefore, the significance level should be larger than 0.95 for the difference to be statistically significant. Accordingly, the p-value must be lower than 0.05 for a difference to be accepted as statistically significant.

From these statistics, we are able to analyze the data that we collect through the experiment and determine the effects of the nudge across genders, groups, and days. In this, we look for immediate and statistically significant differences in means for all the three variables.

DATA ANALYSIS

This chapter centers on the data analysis of the quantitative data gathered through the experiment. It starts with an analysis of EC and AC separately that seeks to determine if there are any observable differences between the three variables; the experimental groups, the four days of the experiment, and the genders, and whether such differences are statistically significant. Afterwards follows an investigation of the potential differences between EC and AC. The differences are identified by creating an Expected Actual Consumption Ratio (EACR), which is presented in the analysis below. The investigation of EACR provides a deeper understanding of the tendencies and differences in the data, and with this we seek to broaden the understanding of any of the observable differences that we discover. The purpose of the analysis is to determine whether the nudge had a different impact on any of the groups, on the genders, or across the different days.

The analysis is based on descriptive statistics (Appendix 5) and comparison of means with the use of independent t-tests (Appendix 6), which determine whether the observable differences in means are statistically significant. With the use of graphs, we visualize the data to be able to better understand the observations (Appendix 7). The primary purpose of the chapter is to determine whether the nudge had a significant effect on the participants' behavior. Second, it seeks to provide a thorough understanding of the different elements that might have impacted the data. In conclusion, the analysis enables us to test the seven hypotheses that were formed prior to the experiment, and the purpose is to accept or reject them based on the findings.

Expected Consumption

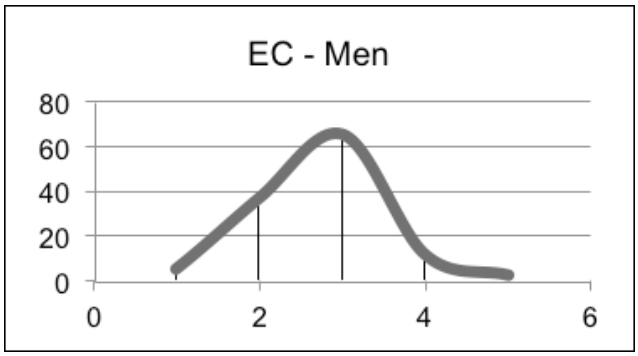
This section looks into the values of expected consumption (EC) gathered from the two groups of both men and women over the period of 4 days. First, the means of EC are descriptively analyzed. Hereafter, we investigate whether the observable differences in means across genders, groups, and days are statistically significant with the use of the independent t-tests.

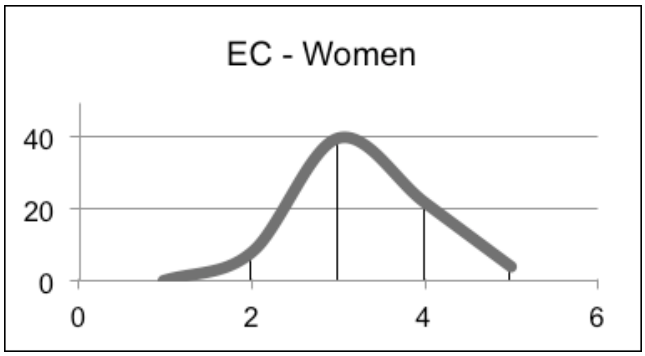
The overall mean for expected consumption across days, groups, and genders is 2.96. This equals almost half a plate of fruit and vegetables. This means that in general, the par-

ticipants expect to consume half a plate of fruit and vegetables for lunch. The mean is derived from 198 observations of expected consumption, collected from both men and women over the period of four days at the two different times of the day. In the following sections, EC is investigated across genders, groups, and days to find out if any observable differences can be ascribed to the three variables.

Genders

Starting out by looking into EC for men versus women, it is found that the women of the experiment in general have higher expectations for consumption of fruit and vegetables than the men. The mean EC for the women is 3.30, i.e. their expected consumption is between ½ and ¾ plate, whereas for the men the mean is 2.76, i.e. that their expected consumption is between ¼ and ½ plate.





From these two graphs, it is seen that the difference in means is not caused by extreme outliers. Instead, the graphs illustrate that the general tendency for the women is that they expect to eat primarily half a plate (3) to ¾ plate (4), whereas for the men, the general tendency is that they expect to eat primarily half a plate (3) to ¼ a plate (2). One might even propose that the graphs are laterally reversed.

Though, this observation might be true, we make an independent t-test to determine whether the difference in means is statistically significant. When comparing EC for the men and the women in the independent t-test, it gives a Sig.-value of 0.947, which means that one can say with a 94.7% confidence that the difference between the genders is statistically significant. This equals a p-value of 0.053. For a p-value to be statistically significant, p must be < 0.05, because in statistics one accepts a difference as statistically signif-

icant if one can say with a 95% confidence that the differences are significant. The p-value 0.053 in this case is very close to the 0.050-line for the difference to be accepted as statistically significant. Therefore we argue that the difference is significant, i.e. in general the women of the experiment expect to eat more fruit and vegetables than the men.

Groups

Hereafter, we seek to determine if there are any differences in EC based on what time of the day the participants are asked about their consumption. For the 9 o'clock group, the mean EC is 3.11 and for the 12 o'clock group, the mean EC is 2.85. From these numbers, one could suggest that the time of day impacts the expected consumption. Though, before making any conclusions on the matter, it is necessary to look into the division of men and women within the two groups, as a higher number of women in the groups will supposedly move the means to a higher number, as the women of the experiment in general expect to eat more fruit and vegetables than the men.

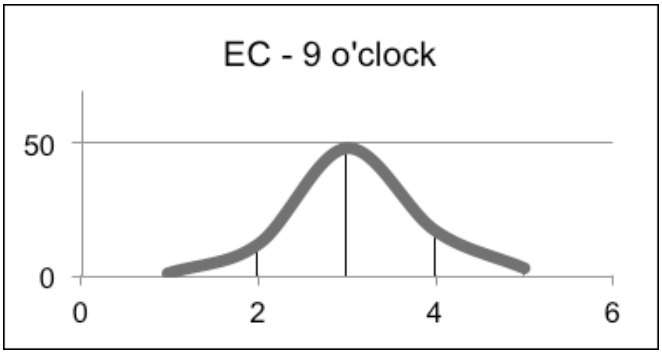
Group	Gender, percentage distribution
9 o'clock	Women = 29, 35%
	Men = 53, 65%
12 o'clock	Women = 46, 40%
	Men = 70, 60%

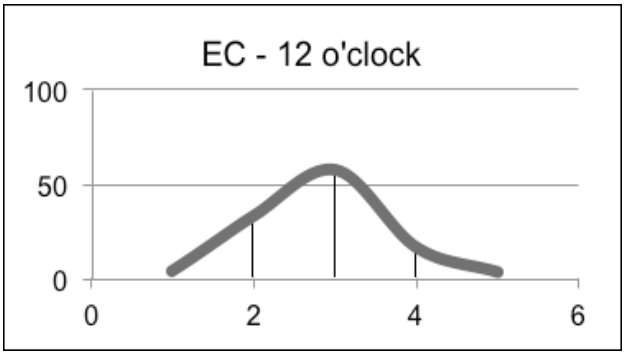
Table 2: Distribution of genders across groups.

The table above shows that the division of men and women differ with 5 percentage points within the two groups, though with more women in the 12 o'clock group, where the mean is lower. Therefore, one might argue that if the division of men and women in both groups had been equal, the difference in the means might have been even larger. To conclude, the difference in means of EC between the two groups could supposedly be a result of the different times of the day.

If looking at the graphs below, it is seen that the difference in means of EC is a result of the 12 o'clock group tending to eat a ¼ plate of fruit and vegetables to a larger extent than the 9 o'clock group. Otherwise, the graphs look similar in distribution of 1 till 5. What is of

interest in this is the question of whether the availability heuristic will then have a distinct impact on the actual consumption as hypothesis H2 proposes.





However, before moving to propositions about actual consumption, we test whether this immediate difference in the means of EC at the two times of the day is even statistically significant. Again, we apply the independent t-test. For this data set, p = 0.84 and therefore the difference in means between the 9 o'clock group and the 12 o'clock group is not statistically significant. This means that the nudge does not have a significantly different impact on the participants' estimates of expected consumption at 9 o'clock versus at 12 o'clock.

Days

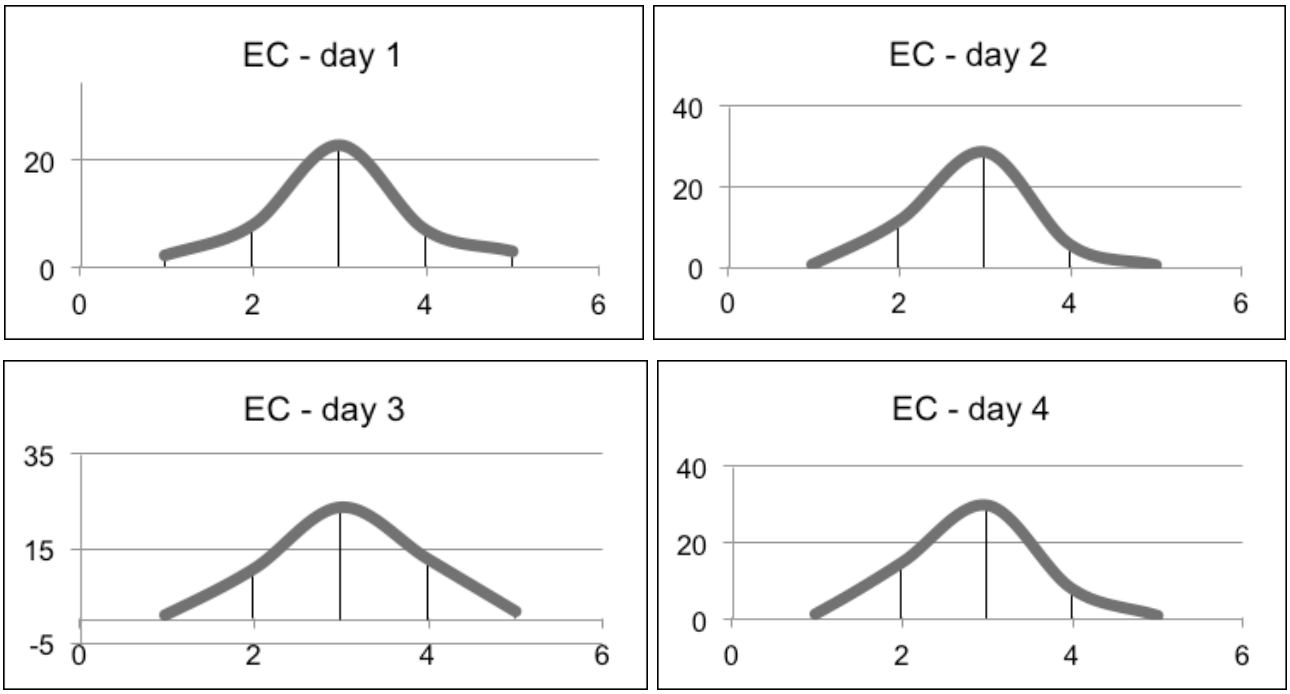
When looking to days, there are quite some immediate differences to observe. Over the four different days, EC is highest on day three, and also higher on day one compared to days two and four. We seek to discover why this applies. If first looking into whether gender can be the cause of difference, it is found that the division of men and women is as follows:

Day	Mean	Gender, percentage distribution
Day 1 (Monday)	3.02	Women = 18, 43%
		Men = 24, 57%
Day 2 (Tuesday)	2.88	Women = 17, 34%
		Men = 33, 66%
Day 3 (Wednesday)	3.08	Women = 17, 33%
		Men = 34, 67%

Day 4 (Thursday)	2.87	Women = 23, 44%
		Men = 32, 56%

Table 3: Distribution of genders across days.

From this, it can be determined that gender is not a relevant contributor to the differences in the expected consumption, as the division of men and women over the four days is almost evenly distributed. Further, the fact that on day 3, the expected consumption (3.08) is the largest and the percentage of women (33%) is the lowest supports the argument that genders are not a relevant contributor to the observable differences. Therefore, the differences must be ascribed to other factors.



Looking at the graphs above, it is found that the expected consumption across the four days is to a certain extent similar, though with slight differences. For example, if looking to day three, the graph is flatter as especially ¾ plate (4) is more frequently observed.

For some reason, there are some observable differences in the expected consumption that is not based on gender. Further, we argue that the differences are also not caused by our presence, as the experiment is run in the exact same way during all four days. Again,

to test whether the assumed differences are statistically significant, we run independent t-tests comparing all four days. The p-values are presented in the table below.

Days	**p-value**
1 vs. 2	0.163
1 vs. 3	0.512
1 vs. 4	0.319
2 vs. 3	0.705
2 vs. 4	0.292
3 vs. 4	0.535

Table 4: p-values for EC across days.

Looking at the six p-values, we can conclude that none of the differences in means are statistically significant. Hence, the specific day does not significantly impact expected consumption of the participants.

To conclude, the only factor that impacts the participants' estimates for EC is gender. The specific day and time of day do not have any statistically significant impact on the expected consumption of the participants. Though, it could be argued that if the sample size had been bigger and if the division of gender had been equal, then more significant differences could supposedly have been determined. However, already at this point, we can conclude that expected consumption is not dependent on the variables groups and days as proposed in hypotheses 6 and 7. On the other hand, it seems that EC is in fact dependent on gender. This will all be discussed further as we test the seven hypotheses.

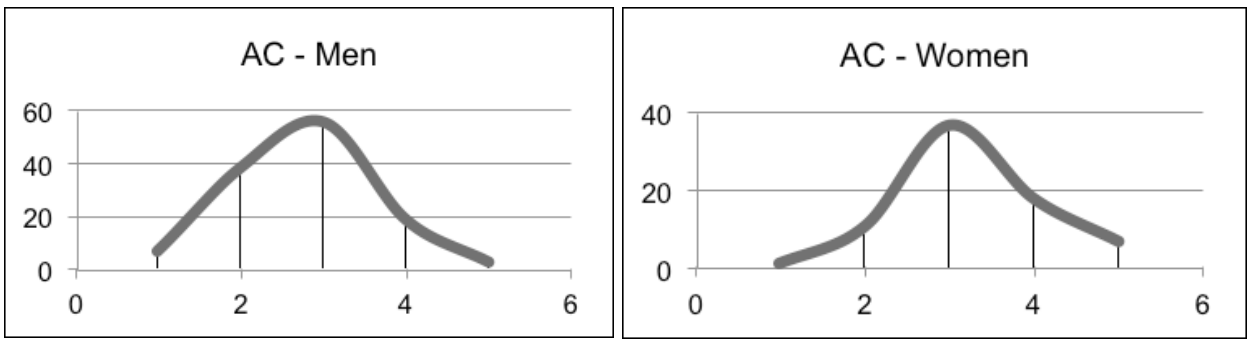
Actual Consumption

In the same way, this section looks into the values of actual consumption (AC) to investigate potential differences in means within the three variables; gender, day, and time of day. The distribution of gender is the same within AC as within EC. Again, the means of AC are descriptively analyzed, and afterwards we determine if the observable differences in means are statistically significant with the use of the independent t-tests.

The overall mean of AC for all participants across days, genders and groups is 2.95, which is very close to half a plate of fruit and vegetables. The mean is derived from the same 198 participants, who were asked to state their expectations for consumption prior to the lunch and hence were nudged. This includes both the participants who were asked about expectations for consumption in the morning and those who were asked right before the lunch break. When looking to the mean AC of 2.95, which is 0.01 point lower than the expected consumption, one might already claim that the nudge did not have the intended effect on the participants' consumption. The overall hypothesis states that the nudge should result in that the participants eat at least the amount of fruit and vegetables as stated before lunch or more. From this immediate observation, the nudge did not succeed to impact the participants as expected. However, we do not reject the hypothesis just yet. First we look into the three variables to determine if there are any observable differences or statistically significant differences caused by genders, time of day or weekday.

Genders

First looking at the difference between genders, the mean AC is higher for the women than for the men. For the men the mean AC is 2.77 and for the women it is 3.26.

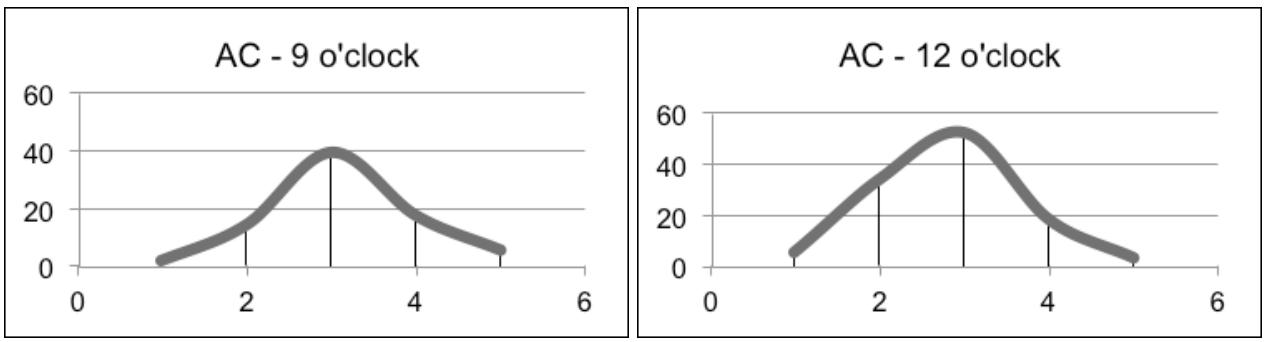


Looking at the two graphs above, it is seen that the difference in means might be a result of a higher observation of 1's for the men and 5's for the women. Such extreme outliers will naturally impact the mean AC. When testing whether this immediate difference is statistically significant, the independent t-test shows that $p = 0.042$, and the difference is therefore statistically significant. With 96% confidence we can say that the women of the experiment eat more fruit and vegetables than the men. This was probably to be expected

as we already observed a statistically significant difference between the genders for expected consumption.

Groups

In this section, we look for differences in AC based on what time of day the participants are asked. It is found that AC for the 9 o'clock group is higher than for the 12 o'clock group. The 9 o'clock group consumed on average 3.14, i.e. between ½ and ¾ plate of fruit and vegetables, and the 12 o'clock group consumed on average 2.83, i.e. between ¼ and ½ plate. Here one could even suggest that because the 9 o'clock group has a smaller percentage of women than the 12 o'clock group, the difference in AC could potentially have been larger if the division between the genders had been the same in both groups.

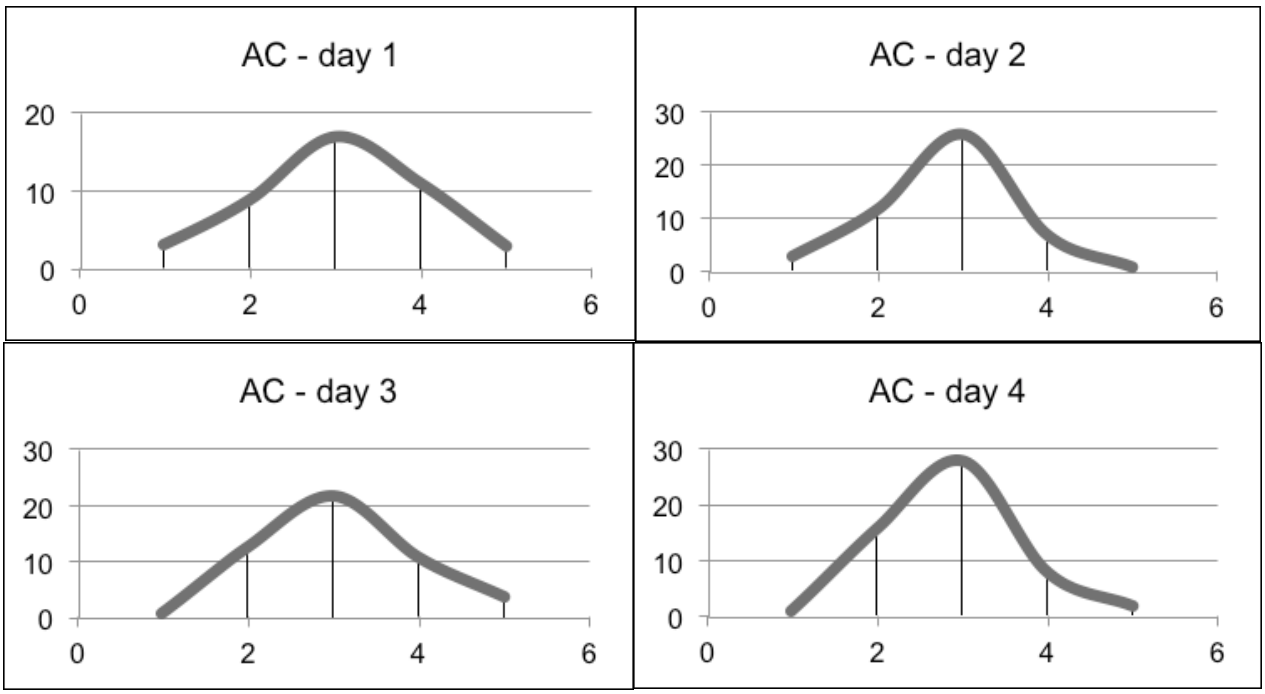


From these graphs, it seems evident that the difference in means between the two groups is primarily caused by the more frequent observation of ¼ plate (2) in the 12 o'clock group. This tendency is similar to what we observed already within the expected consumption. Therefore, one could suggest that the participants in fact do as expected, and hence that the availability heuristic did not make any noticeable impact. We look further into whether this is the case when looking into the expected actual consumption ratio.

To test whether this immediate difference is statistically significant, we apply the independent t-test, which gives a p-value of 0.222. Evidently, the difference in means is not statistically significant, and therefore the immediate differences between group 1 and group 2 and this higher observation of ¼ plate in the 12 o'clock group could be caused by mere chance.

Days

Finally, looking into whether there are differences in AC across the four days, it is found that AC was largest on day 1 and that day 2 had the lowest average consumption. Again, as the split between genders differs a bit across the four days, it might have a small impact on the results but as with EC this is not considered to have a critical impact.



The independent t-tests give the following p-values:

Days	**p-value**
1 vs. 2	0.719
1 vs. 3	0.356
1 vs. 4	0.836
2 vs. 3	0.489
2 vs. 4	0.255
3 vs. 4	0.675

Table 5: p-values for AC across days.

Again, none of the differences in means are statistically significant, which means that the specific day does not have a significant impact on the actual consumption. As we also

found within EC, the specific day does not influence the actual consumption of the participants.

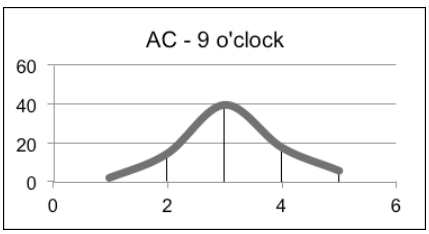
To sum up, we observed quite some immediate differences within the actual consumption of the participants across genders, days, and groups. However, the only statistically significant difference within AC is the different consumption between the genders, like we observed for expected consumption. One could then suppose that the nudge did in fact not have any significant impact, and that the difference between genders would also appear without the nudge. We investigate whether this is the case by bringing in the control group.

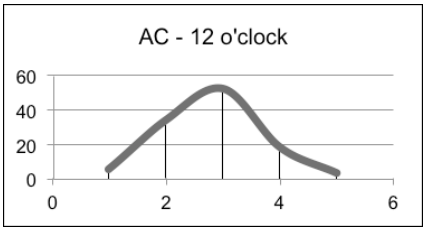
Control group

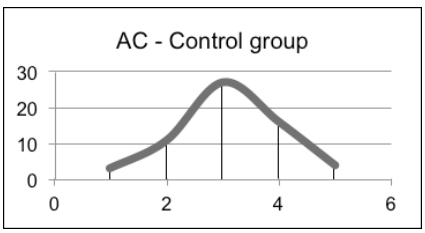
We have included the control group with the purpose of determining the mean actual consumption without the presence of the nudge. The mean AC of the control group is 3.11. This indicates that without influences of any kind, an average user of the canteen would eat between ½ and ¾ plate of fruit and vegetables. In this section, we look into the difference in AC between the control group and the two experimental groups, to determine if the potential differences are significant. Additionally, we can use the control group to determine if this difference that we have observed between the genders does also apply within the control group.

Groups

When comparing the results of the control group with the experimental groups, we see that the 9 o'clock group eats around the average (3.14), whereas the 12 o'clock group eats less than the average (2.83). This could indicate that the nudge did actually have a negative impact on the participants of the 12 o'clock group.







To verify whether any of these differences are statistically significant, we use the independent t-test. This gives the following results:

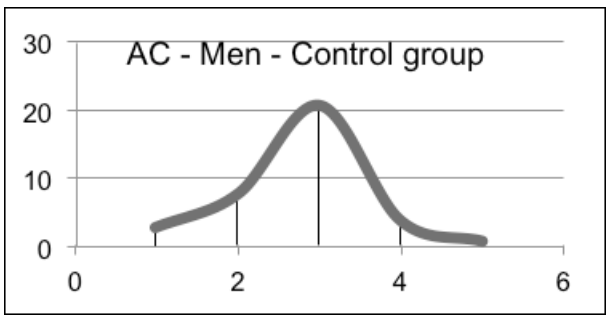
Groups	p-value
1 vs. 3	0.382
2 vs. 3	0.24

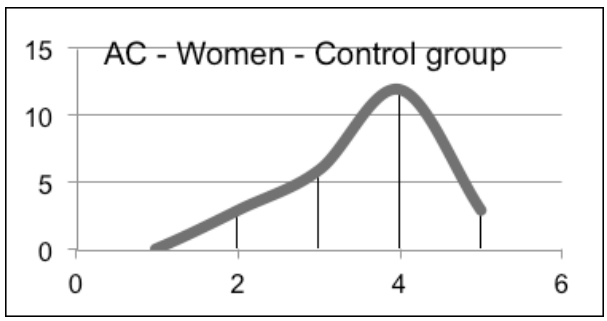
Table 6: p-values for AC across groups

Evidently, the differences in means between all three groups are not statistically significant. This indicates that the nudge did in fact not have the intended effect on either the 9 o'clock group or 12 o'clock group. The 9 o'clock group eats about the same amount of fruit and vegetables as the control group and the 12 o'clock group eats even less fruit and vegetables than the control group. This indicates that the nudge might have had a negative effect on consumption behavior. Reasons for why this may be are presented in the discussion.

Days

Another element that we find interesting to investigate is whether the statistically significant difference between the genders is also evident within the control group. Within the control group, we find as well an observable difference in the means of AC between the men and the women. The mean AC for the men is 2.78, i.e. they eat between ¼ and a ½ plate of fruit and vegetables, and for the women, the mean AC is 3.63, i.e. they eat between ½ and ¾ plate of fruit and vegetables.





From these graphs, we find that the women of the control group have eaten primarily ¾ plate (4). During the experiment, this is the only time that we observe ¾ as the most frequent value. To test whether the immediate difference between the genders is statistically significant, we again apply the independent t-test. This gives a p-value of 0.479, i.e. the difference between the genders in the control group is not statistically significant. This result then tells us that the differences found between the genders within EC and AC for the two experimental groups are not a result of a general difference between the genders. If this was the case, we would observe the same statistically significance when the participants are not nudged, which is the case of the control group. Therefore we suggest that the statistically significant differences in both EC and AC between the men and the women are in fact an outcome of the nudge. This finding will be further developed in the discussion.

Expected Consumption vs. Actual Consumption

If looking into the standard deviations of EC and AC, it is found that there is in general a bigger variance in the observations of AC compared to EC across all the different variables (gender, groups, and days). If first looking to EC, the different standard deviations for all variables are found within quite the same range; from 0.726 to 0.913, i.e. the observations vary with around the same level across genders, groups, and days. This is also indicated in the graphs above, where we find that the mean values are not a result of extreme outliers but instead a result of frequent observations of primarily ½ a plate, and to some extent ¼ and ¾ plate. All observations found within +/- 1 standard deviation include 68% of the observations. In general, this means that 68% of our observations are found within the range of around ¼ plate and around ¾ plate, as most of the observed means are very close to the value of ½ a plate. When looking to AC, the variance is a bit larger and the standard deviation across all variables ranges from 0.809 to 0.977. This tendency suggests that the participants have more similar ideas about what they expect to consume of fruit and vegetables but when looking to what actually happened, the differences in the observations are in general larger.

From the first part of the analysis, it is found that the only significant factor in the results is the difference between the genders. The difference in means between the genders is sta-

tistically significant both within expected consumption and actual consumption. Therefore, one could suggest that the differences are not a result of the nudge but instead the result of basic differences between the genders in terms of eating fruit and vegetables. However, the results of the control group show that this is not the case and that the observable difference between the genders of the control group is not statistically significant. Therefore, we propose that the nudge proves to have a different effect on the genders.

Even though the results of the groups and days show that there are no statistically significant differences in the consumption of the participants, we still find that there are some interesting differences in behavior that we seek to investigate in further detail. Additionally, it is considered that most of the differences might be insignificant because of the small sample sizes. The greater the sample sizes, the bigger the likelihood of being able to determine statistically significant differences. Thereby, it is argued that if the experiment had had more observations we might have observed more significant results. Therefore, we take a deeper look at the differences between EC and AC below to get a deeper understanding of the psychological elements at play.

Expected Annual Consumption Ratio (EACR)

The next section goes into a deeper analysis of the differences between expected consumption and actual consumption, and therefore an expected actual consumption ratio is calculated. We call it EACR and it is the difference between AC and EC, i.e. AC minus EC. When EACR is negative the actual consumption is lower than expected and when EACR is positive, the actual consumption is higher than expected. This indicates that when EACR is negative the consumption was overestimated, whereas when EACR is positive the consumption was either underestimated or the nudge had a positive effect on the consumption behavior of the participants.

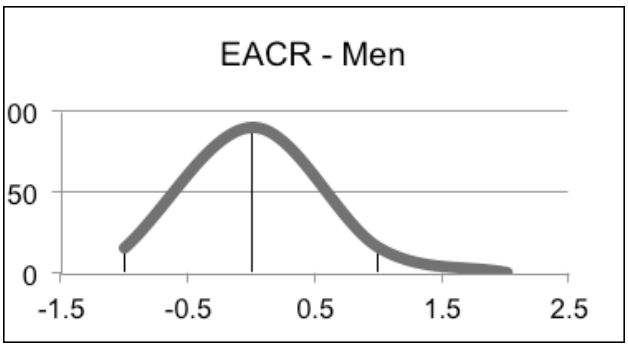
Across all groups, days, and genders, the mean of EACR is -0.005. This number indicates that in general the participants of the experiment have a slight tendency to overestimate in their expectations for consumption. However, the small number indicates that there are both observations of overestimation and underestimation, which even out one another. It is found that 72% of the participants eat exactly the amount of fruit and vegetables as they

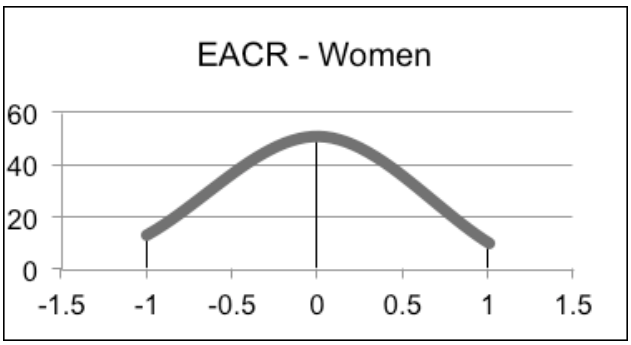
expect, and only 28% eat either more or less than expected. The division between eating less versus eating more is 13% and 15% respectively, which again shows that the general tendency to overestimate is almost non-existing and that to almost the same extent the participants tend to underestimate.

In the following, we look into EACR for genders, groups, and days and investigate whether there are any significant or observable differences in EACR across the variables.

Genders

The mean EACR for the women is -0.04 and 0.02 for the men. This indicates that in general the women have a tendency to overestimate their consumption, whereas the men either underestimate their consumption or they are impacted by the nudge to eat more than expected. From the graphs below, we find that the difference in means is not caused by an overweight of misestimates or many extreme outliers (we only observe a single value of '2' within EACR of the men), but that it is primarily a result of the frequent observations of 0's. However, when applying the independent t-test on the results, it gives a p-value of 0.557 and the difference in means is therefore not statistical significant. Thereby it cannot be assumed that the women will typically overestimate and the men will typically underestimate. Further, we cannot say whether the nudge had a significant impact on the men and not the women, as the observable difference might be found by chance.

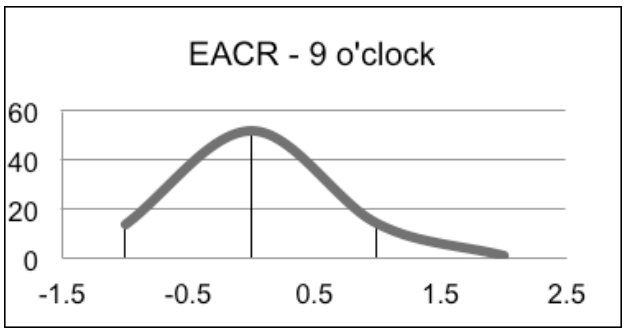


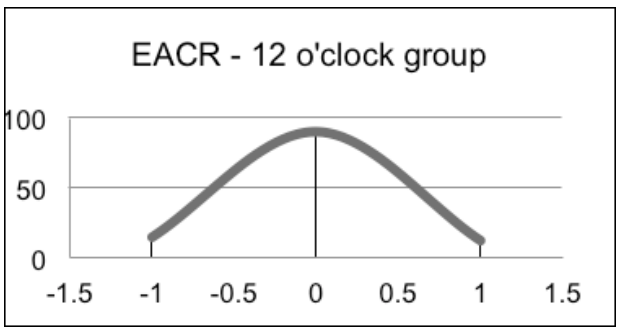


Groups

For the two groups, it is found that the 9 o'clock group has a tendency to underestimate with an EACR of 0.02, whereas the 12 o'clock group overestimates with an EACR of -0.03. Again, the results are not caused by extreme outliers but by the frequent observation of 0's

and a range of -1's and 1's. However, the difference between the 9 o'clock group and the 12 o'clock group is not statistically significant as p = 0.963. From this result, there is no indication that the nudge had a different effect on the two groups.





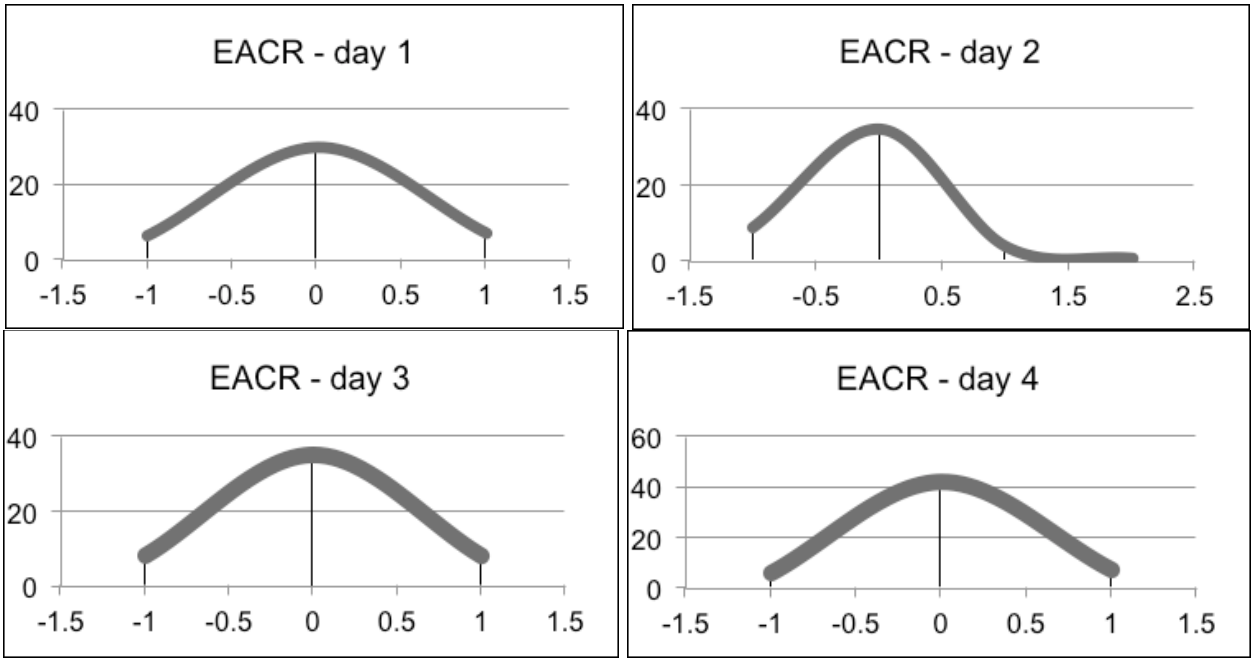
Days

When looking at the differences between the days, we again observe some immediate differences in means:

Day	**Mean**
Day 1 (Monday)	0.02
Day 2 (Tuesday)	-0.06
Day 3 (Wednesday)	0.00
Day 4 (Thursday)	0.02

Table 7: Mean Expected Annual Consumption Ratio across days.

Monday, the participants underestimate consumption and EACR is 0.02, whereas on Wednesday the participants eat exactly what they expect with EACR being equal to 0.00. Tuesday, the participants overestimate with an EACR of -0.06 and Thursday the participants again underestimate with EACR of 0.02. From the graphs below, we can observe some small differences in the frequency of the observations of -1, 0, 1, and 2, but in general the graphs look rather similar, with Tuesday being the only exception with a single observation of the value '2'.



Independent t-tests are applied to test whether the differences in means of EACR are statistically significant. It is found that between Monday (day 1) and Wednesday (day 3) the difference in the results is statistically significant with a p-value of 0.036. The differences across the others days are not statistically significant and they are therefore not further commented.

Days	**p-value**
1 vs. 2	0.216
1 vs. 3	0.036
1 vs. 4	0.556
2 vs. 3	0.257
2 vs. 4	0.716
3 vs. 4	0.538

Table 8: p-values for EACR across days.

To conclude any further on why there is a significant difference in EACR between Monday and Wednesday, one should conduct a similar experiment over more weeks to uncover whether such difference would repeat itself and to be able to determine what could be the cause of the difference. Evidently, the day proves to have a different impact on the effect of the nudge on Monday than on Wednesday, which means that the days to some extent

moderate the nudge, as we also propose in the hypothesis development. This is further commented as we test the hypotheses.

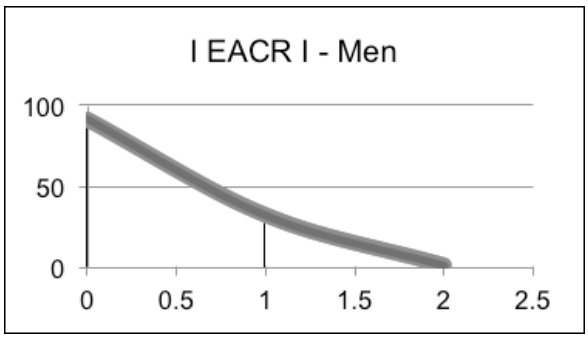
| EACR |

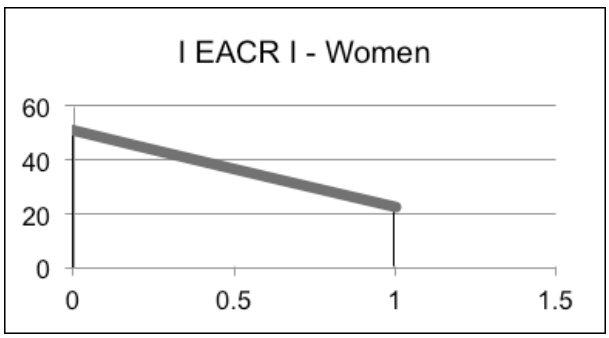
This section looks into the numerical value of EACR with the purpose of determining the degree to which the participants of the experiment are not consistent with their expectation, i.e. to what extent EC does not equal AC. The numerical value of EACR enables a better understanding of the actual misalignment between EC and AC, where a negative EACR-value is not equalized by a positive EACR-value. The numerical results do therefore not indicate whether the participants overestimate or underestimate their consumption.

The mean of | EACR | across days, gender, and groups is 0.29. This means that on average people misestimate their consumption by 0.29. In the following we look into potential differences within | EACR | across genders, groups, and days.

Genders

Looking to the genders once again, we find that the inconsistency between expected and actual consumption is larger for the women than for the men. The mean | EACR | for the women is 0.31, whereas for the men it is 0.27.



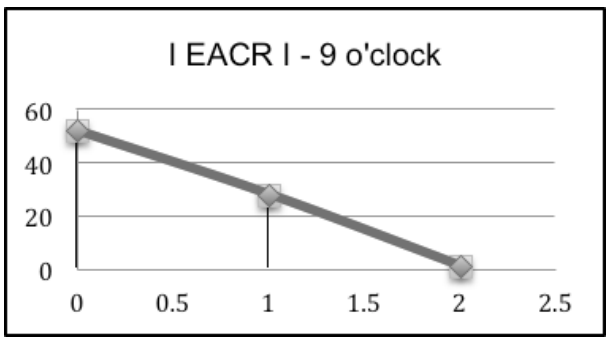


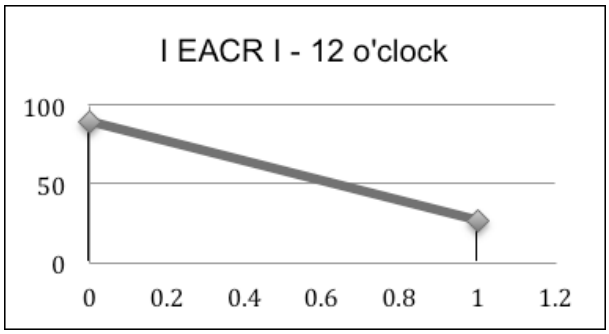
From the graphs above, it is seen that the actual consumption differs from the expected consumption for around a fourth of the men, whereas for the women it is almost a third, for whom the actual consumption differs from their expectations. Again, the difference in means is not caused by extreme outliers, but it is a result of a single increase or decrease

in consumption (with only one exception). Though, the difference in means is not statistically significant as the result of the independent t-test gives a p-value of 0.593.

Groups

Between the 9 o'clock group and the 12 o'clock group, we find that there is a larger inconsistency between EC and AC within the 9 o'clock group. The mean |EACR| for the 9 o'clock group is 0.37 and for the 12 o'clock group is 0.23.





From these graphs, it is seen that the inconsistency between EC and AC is largest for the 9 o'clock group where more than a third of the participants eat either more or less than they expect to do. For the 12 o'clock group, it is less than 25% of the participants that misestimate their consumption. With a p-value of 1, the difference is though not statistically significant. Therefore, we cannot conclude that the nudge impacted the groups differently even though the immediate results might indicate so.

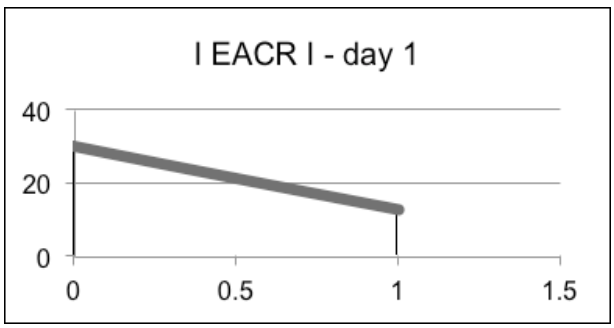
Days

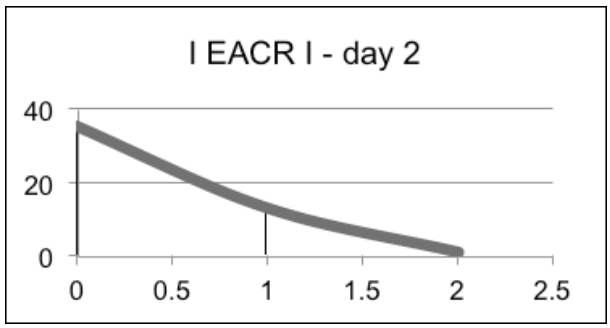
For days, the mean |EACR| is as follows:

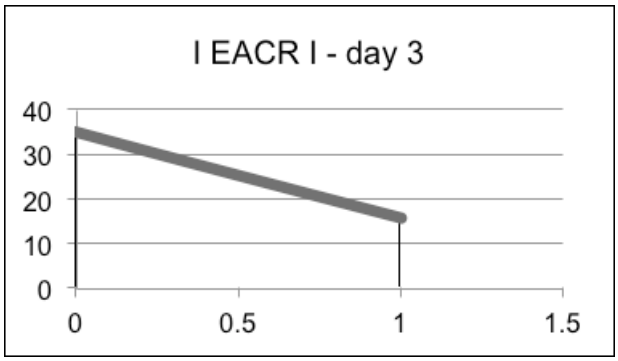
Day	**Mean**
Day 1 (Monday)	0.02
Day 2 (Tuesday)	-0.06
Day 3 (Wednesday)	0.00
Day 4 (Thursday)	0.02

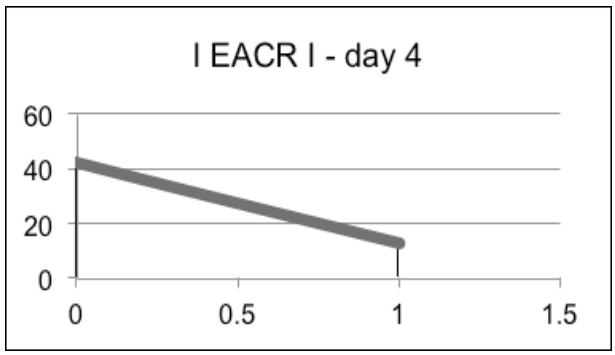
Table 9: Mean |EACR| across days.

From the means one could suggest that there is not much difference to observe across the four different days. Only day 4 seems to differ a little from the others. If looking to the graphs below, one could suggest that day 2 differs from the others, even though the mean is similar to the other days. However, none of the differences in means are statistically significant when tested with the independent t-test and it can be concluded that the specific day does not impact the differences between the expected consumption and the actual consumption.









Interestingly, the statistically significant difference that we observe between Monday and Wednesday within EACR is not observed with the numerical value. This again underlines why more investigation should go into a similar experiment conducted over a period of several weeks to understand why such difference between Monday and Wednesday is observed.

Days	**p-value**
1 vs. 2	0.260
1 vs. 3	0.186
1 vs. 4	0.845

2 vs. 3	0.123
2 vs. 4	0.894
3 vs. 4	0.919

Table 10: p-values for ∣EACR∣ across days.

After looking into EACR and ∣EACR∣, we find that the only statistically significant difference that we observe is the difference in means of EACR between Monday and Wednesday. At this moment, we find that there is no logical explanation for why this difference applies, which is why we propose that further research should go into this element to determine what might be the cause of it. Further, the statistically significant difference between the genders, which we identified within both EC and AC, do not apply for EACR and ∣EACR∣. Therefore, we find that the genders react differently to the nudge in terms of their statements of consumption but their behavior from expected consumption to actual consumption is rather similar, or at least not statistically significantly different. In other words, the genders are different in terms of statements and commitments of expected consumption but both genders do to a large extent live up to their individual commitments. Therefore, we can conclude that the genders respond similarly to the availability heuristic and we propose that other biases or heuristics impact their different responses about their expectations.

Hypothesis Testing

This section rounds off the analysis with testing the seven hypotheses. In this, we seek to accept or reject the hypotheses that we developed prior to the experiment. We start out by testing the six sub-hypotheses, which we derived from the overall hypothesis. The overall hypothesis is then tested and finalizes the analysis.

First of all, we can in general terms not accept any of the hypotheses, which propose that there are differences to be found across days and groups, as none of the immediate differences proved to be statistically significant. However, we still look into the individual hypotheses and test them based on the immediate differences that we observed. We start out by looking to hypothesis 2:

H2: The nudge will have a larger effect on group 2, who has most recently stated the expectation for consumption, than on group 1, i.e. there will be a larger increase in consumption for group 2 than for group 1.

The first sub-hypothesis, H2, is rejected. The immediate difference between the two groups shows that the results are actually reversed. When looking to EACR we observe an overall increase in actual consumption for the 9 o'clock group, whereas for the 12 o'clock group we observe a decrease in consumption when compared to their expectations. Though as the difference in EACR between the groups is not statistically significant, we cannot be sure whether the observable difference is caused by coincidence, and therefore we cannot either accept the null-hypothesis, i.e. the nudge will have a larger effect on group 1.

The second sub-hypothesis suggests that there is a difference in the consumption between men and women.

H3: Women will be more affected by the nudge than men, i.e. the increase in consumption will be greater for women than for men.

We did observe a statistically significant difference between the genders both within EC and AC, and we found that the difference was a result of the nudge in one way or the other. However, the hypothesis must be rejected, as we did not observe a greater increase in consumption for the women than for the men, in fact the opposite. The results show that women tend to overestimate, which is why we observed a decrease in the actual consumption compared to the expectations of the women. However, these observations are not statistically significant. By rejecting this hypothesis, we should instead be able to accept the null-hypothesis, which is that women are less affected by the nudge than men. However, as the difference is not statistically significant, we cannot say with certainty say that the null-hypothesis is true. Whether the null-hypothesis might in fact be true is further elaborated in the discussion.

Hereafter, we look to the two hypotheses that propose that differences caused by the commitment heuristic can be observed. For H4 we find that this hypothesis is close to the

truth, as a larger percentage of the participants of the 12 o'clock group eat exactly what they expect, compared to the 9 o'clock group. Evidently, the difference between EC and AC is larger for the 9 o'clock group than for the 12 o'clock group. This is found when looking to |EACR|. The mean |EACR| for the 9 o'clock group is 0.37, whereas for the 12 o'clock group it is only 0.23. The inconsistency between EC and AC is hence largest for the 9 o'clock group as H4 suggests. However, the difference in means is not statistically significant. Therefore, we reject H4 on the basis that we cannot know whether this difference is observed by chance. Second, H5 is as well rejected, as we observed no statistically significant differences between the misestimation and herein the commitment across the four days.

> *H4: People are more inclined to keep a recent promise of consumption (group 2) than an earlier promise (group 1), i.e. the difference between EC and AC will be larger for group 1.*

> *H5: People are more inclined to keep their promise of consumption in the beginning of the week, i.e. EC is closer to AC in the beginning of the week*.

The next two hypotheses set forth to test the same differences with regards to optimism.

> *H6: People overestimate healthy consumption in the beginning of the day, i.e. the difference between EC and AC will be larger for group 1 than for group 2.*

> *H7: People overestimate healthy consumption in the beginning of the week, i.e. the difference between EC and AC will be larger in the beginning of the week.*

First, by looking at the numbers for EACR, it is found that H6 should be rejected as the 9 o'clock group in fact underestimates their consumption. On average, the 9 o'clock group eats more fruit and vegetables than expected, whereas the 12 o'clock group eats less. Yet again, as the difference in means is not statistically significant we cannot accept the hypothesis. Second, looking at the different days we find that there is no logic coherence be-

tween the immediate differences across days, even though there are quite some observable differences between the four days. Generally, the participants tend to underestimate on Monday, whereas on Tuesday they overestimate. This is found by looking to EACR. Monday and Tuesday are considered the beginning of the week but as we observe that on Monday the participants underestimate their consumption we cannot accept the hypothesis. However, within this part we observe that the difference in EACR between Monday and Wednesday is statistically significant, which indicates that there might in fact be some kind of difference to observe in terms of optimism across the different days of the week. Though, as this is not in line with what the hypothesis proposes we still cannot accept H7.

Finally, we test the overall hypothesis:

> *H1: Participants stating their expectations for consumption behavior on consumption of fruit and vegetables will result in at least the stated consumption or an increased consumption of fruit and vegetables.*

From the results, it is evident that we did not succeed to nudge the participants to increase their consumption on a general basis. Looking to EACR, we found that the overall mean EACR is -0.005, which indicates that there is a slight tendency to overestimate consumption among the participants. However, this very small difference might just as well be caused by chance. However, we could argue that because the mean EACR is so close to zero and because 72% of all the participants eat the exact amount of fruit and vegetables as stated, the nudge proved successful to some extent in terms of getting the participants to eat at least the stated consumption. Despite these findings, we will still reject the hypothesis, as 72 % of the participants is not significant enough to accept the hypothesis.

Sub-conclusion

During the analysis, we observed a range of immediate differences and tendencies within the data that we collected through the experiment, which we propose can be an outcome of the intervention that we inserted into the decision-making situation. With the experiment, we sought to determine whether stating an expectation for consumption could function as a recent memory that could be recollected in the moment of decision-making and hence impact the decision. From the analysis, it is evident that the statement for expected con-

sumption did in general not succeed to increase the consumption of the participants. Instead, we observed some other interesting tendencies, the primary being the statistically significant difference between the genders in EC and AC. We find that there is a statistically significant difference between the genders in terms of both their expectations for consumption and the actual consumption. Though, what is interesting is the fact that this difference does not apply in the control group, and hence, the difference is not an outcome of a natural difference between the genders. This indicates that the nudge had a different effect on the men and the women in the element of stating their expected and actual consumption. As no statistically significant differences are found in either EACR or |EACR| for the genders we can, at this moment, not come any closer to an explanation for why this difference applies but we will expand on this difference in the discussion to obtain a better understanding of the dynamics at play. However, one thing that we can say is that the differences between the genders are not caused by the availability heuristic. Further, as all the other immediate differences are not statistically significant, we cannot claim that the availability heuristic proved as the vital or only contributor to the results.

Finally, we propose that an even distribution of women and men amongst the participants and across the two groups and the four days would have been more appropriate and might give a more rightful result, exactly as the only statistically significant difference that we identify is between the genders.

DISCUSSION

This chapter seeks to provide a deeper understanding of how the nudge sought to impact the participants of the experiment. Through discussions of both quantitative and qualitative observations, we seek to broaden the investigation of the effect of the nudge. This leads us onto suggestions for future research, and finally it enables us to make a conclusion on the research. The chapter starts out with a post-talk analysis that presents a range of qualitative observation that we made during the experiment. These observations are found to be important aspects to consider when evaluating the actual outcome of the experiment. The observations are used as a basis for further discussion at a later stage in this chapter. Second, we move onto the results of the analysis, where we seek to broaden the understanding of whether or not the nudge proved successful. In this, we give special attention to the statistically significant difference found between the genders. Afterwards follows further discussions of the observations and considerations that emerged out of the experiment. As a final part of the discussion, we give attention to ethical considerations about nudging and the general implications that follow from applying nudging as a marketing and communications tool.

Post-talk Analysis

This section covers observations that we made during the experiment. The observations are a combination of pure observations of the participants and statements and comments made by the participants. During the experiment, we noted the observations on a piece of paper at the exact moment that they happened for them to be as objective as possible. By following this method, we ensure that the observations are as accurate as possible, as we note them when they are fresh in mind and hence not blurred or biased. The observations do not impact the results of the analysis directly but they are important aspects to consider to get a deeper understanding of the specific problem area. Further, they help broaden the understanding of how the different heuristics and biases come into function at different stages of the decision-making process.

The post-talk analysis starts with some general observations that we made during the experiment. Following this, we present some methodical observations with the purpose of

understanding how the experimental setup might have impacted the results. This enables a discussion about how to further investigate the problem area and which methodical considerations to take into account. The section ends with a presentation of a set of limitations that we discovered during the experiment. The purpose of the post-talk analysis is to take the quantitative data and enrich them with a set of qualitative observations, which gives a deeper understanding of the subject matter. Herein, we combine the quantitative study with a set of qualitative data and we mix the two methods to strengthen the research.

General observations

Fruit and vegetables is a highly debated topic and during the days of the experiment it became evident that the participants often times supposed that we were investigating matters of health, and that the experiment sought to understand whether the participants were healthy or not. On a general basis, we sought to inform the participants that we were not investigating matters of health but that we simply wanted to understand consumption habits, and therefore we sought only honest answers.

It became clear that most of the participants had a clear understanding of the amount of vegetables and fruit that they normally eat for lunch. Almost none of the participants had any difficulties defining how much vegetables and fruit they put on their plates and they were very conscious about their choices. Though, often times the participants' responses were expressed in ranges, stated as: "*somewhere in between plate 2 and 3*" (referring to the pictures of the plates) and the participants were of the conviction that their consumption depended on selection, i.e. the participants knew that the consumption of fruit and vegetables would normally range from e.g. ¼ to ½ depending on the selection of the day. Clearly, the participants were very much aware of their choices. This indicates that eating fruit and vegetables is a very inveterate habit and that the participants follow the same pattern every day. Therefore, it is argued that eating fruit and vegetables is a deliberate and habitual action, which is learned through repeated practice and rewards in the stable environment of the canteen. This might impact which of the participants' brain-systems that are activated during the moment of decision-making, both in terms of stating an expectation for consumption and when making the actual choice by the buffet about the amount of fruit and vegetables. One could suppose that the act by the buffet normally activates the habit-

ual system but with the nudge clear in mind, the impulsive system or the goal-directed system might be independently or jointly activated. This will be further discussed at a later stage.

As mentioned, the participants were of the conviction that the selection in the canteen influences their responses for both EC and AC. The topic was brought up several times by the participants both before and after their lunch break. Before going to lunch, some of the participants mentioned that their response depended on the selection at the buffet. Monday (day 1), the dish of day was Spaghetti Bolognese, which was a very popular dish among the participants. Many of the participants mentioned that they ate less vegetables and fruit than stated before going to lunch because of the dish of the day. During the other three days, the quality of dish of the day was brought up several times and according to the participants, it greatly impacted their consumption of fruit and vegetables.

However, when looking to the analysis, it is found that this conviction does not correspond with the data that we have collected. First of all, the value that is observed most often for AC across all days is ½ a plate (picture 3). This indicates that the selection does not have such a great impact on the amount of fruit and vegetables consumed as the participants suppose. Though, it must be mentioned that we do not include the same set of participants across the four days but we ask a new set of participants every day. This means that we cannot know with certainty that the selection will not influence the consumed amount of fruit and vegetables for the individual participant, but in general terms, this seems not to be the case. Second, in the Expected Actual Consumption Ratio, it is found that Monday the participants actually tend to underestimate their consumption of fruit and vegetables with EACR being positive, even though many of the participants argue that they might have eaten less because of the Spaghetti Bolognese. Further, looking into $|EACR|$ for the different days, we find that the misalignment between EC and AC is smaller on Monday than on Tuesday and Wednesday, i.e. on Monday the participants are better at predicting their actual consumption. Yet again it is found that there is no statistically significant difference in means of EC, AC, EACR or $|EACR|$ across the four days (besides from the difference in EACR between Monday and Wednesday). Therefore one could suggest that the selection in the canteen might not have such a great impact on consumption of fruit and vege-

tables after all. This again supports the argument that the consumption of fruit and vegetables is in fact largely driven by habit, and no matter the selection in the canteen, one tend to eat around the same amount every day.

Taking into account that health and eating fruit and vegetables is such a highly debated topic, it might be a sensitive topic to disclose for certain people. One could suggest that healthy eating is especially sensitive for women, which we already proposed in the hypothesis development. During the experiment, more of the women spent some time explaining why they did as they did and why they did e.g. not take as much vegetables as they claim that they normally do. On the other hand, most of the men were more immediate and gave their response without further explanation. It would be interesting to investigate whether this is a general tendency, or whether we would observe another dynamic if the topic at hand was different. If one were to choose another subject such as beer or wine, it could be suggested that the situation would look differently or even be reversed in matters of sensitivity for men and women. We argue that these observations support the idea that men and women react differently to the nudge and that this might help explain the differences in the means of EC and AC between the genders.

Methodological considerations and limitations

The experimental design and the fact that the experiment is conducted in the field and not in a lab naturally cause some risks, which can impact the results of the experiment. In the following, we present some general methodological observations that might influence the data.

The employees at the ISS Tower and the participants of the experiment generally showed an interest in the research and the experiment itself. It was clear that the subject was discussed in the different offices. This might add a weakness to the results of the experiment, as some of the participants possibly already knew of the scope of the experiment from their colleagues before participating themselves. Though, still on Thursday (day 4) a lot of the employees did not have any knowledge of the experiment when being asked to participate

In some instances, the participants came in groups and they were collectively asked about both expected consumption and actual consumption. This could impact the participants' individual responses, as they might be keen to give the same or similar answers as the rest of the group. This is yet another bias that can come into function and impact the decision-making. Though, it is of our conviction that most people gave their answer independently without being affected by co-workers' or group members' answers. Further, we observed that more of the participants, who were asked in groups, gave different answers. We choose not to go further into a discussion about group mentality and the tendency to act as the group but we point to the fact that another bias should be added to the list of potential heuristics that can impact decision-making. Hence, as already proposed in the methodology chapter, the list of heuristics and biases that we attend to might not be the full list of biases that come into action.

As already explained, to ensure that we were able to keep track of the participants and minimize the dropout, we found it necessary to let the participants know that we wanted them to come back to us after their lunch for a follow-up question. Some of the participants, especially the ones in the 12 o'clock group, already guessed that we would ask them about their actual consumption after the lunch break. Having such supposition inevitably impacts the behavior of the participants. This limits the understanding of whether the availability heuristic had an effect on the behavior or whether a change in behavior was merely a result of the fact that the participants knew that we were going to hold them accountable for their answers. In this, we cannot know whether the potential change in behavior is caused by the availability heuristic, the commitment heuristics or even other heuristics and biases.

Further, there is a possibility that the participants of the 9 o'clock group were reminded of their statements about expected consumption right before their lunch break as they passed by us before entering the canteen. This might affect the power of the nudge, as these participants were supposedly impacted by the nudge twice.

Choosing to use pictures to show the division of the plate and the five possible choices (0, ¼, ½, ¾ and 1/1) was a successful way to get the participants to estimate their consumption. We first considered having the participants estimate their consumption in terms of

how many grams they would expect to consume. Such an estimation would have been very difficult for the participants to make and almost impossible for us to verify both for EC and AC. With the visual element of the plates, the majority of the participants found it rather easy to decide on both EC and AC by just pointing at a picture. Furthermore, this visual element enhances the degree to which we can trust the participants' responses, since the pictured plate from the canteen is recognizable for the participants and is shown in the real environment and in well-known surroundings.

Though, the choice of dividing the plate into four parts proved successful, we found that many of the participants would define their consumption of vegetables and fruit right between the ¼ plate and ½ plate or between the ½ plate and the ¾ plate. One could suggest that we could have benefitted from dividing the plate into only three parts instead of four and thereby allowing the participants to define their consumption more precisely. However, such division would give fewer possible answers and supposedly result in that other participants would find themselves being right in between ⅓ and ⅔. Especially as the results show that the ½ plate is the most frequently observed value during all four days and across groups and gender (the women in the control group being the only exception). Though, dividing the plate into three parts corresponds better with the Y-plate of the Danish Veterinary and Food Administration. The Y-plate recommends eating 2/5 plate of vegetables and fruit during all meals. One could suggest that the official recommendation from the Danish Veterinary and Food Administration might be the cause of the frequent observation of ½ a plate, as this general conviction that between ⅓ and ⅔ a plate of fruit and vegetables is equal to being healthy. In this sense, the choice of consuming around half a plate of fruit and vegetables seems to be a deliberate choice. One could also propose that the Y-plate functions as an anchor, and that the anchor is a starting point for the habit.

Sub-conclusion

Combining the quantitative results of the experiment with a set of qualitative observations made through the experiment enables a more thorough understanding of the elements at play when seeking to impact the behavior of the participants. Only from such knowledge, we will be able to determine if and why the nudge had an effect. The post-talk analysis indicates that the consumption of fruit and vegetables for most people is driven by habit and

to some extent impacted by social norms and tendencies in society. Further, many of the observations indicate that the various heuristics and biases impact the decision-making process at several stages and with varying degree. Such observation must be taken into account when trying to understand how the nudge can be used to impact decision-making.

Discussion of findings

In this section we take a thorough look into the findings that we made through the data analysis of the data collected through the experiment. The discussion brings in elements from both the quantitative data analysis and from the findings that we made in the post-talk analysis. This enables a deeper understanding of the observations and differences that we have identified, and with this we seek to come closer to an identification and isolation of the availability heuristic.

Genders

The main finding of the analysis is that the nudge seemed to impact the genders differently. Both the expectations and the actual behavior of the men and the women in the experimental groups differ in terms of eating fruit and vegetables. These differences are statistically significant. As the immediate difference between genders of the control group are not statistically significant, it is suggested that the men and the women of the experimental groups react differently to the nudge. If the differences are to be caused simply by natural differences between genders, one should be able to observe the same difference within the control group but since this is not the case, we find that the nudge succeeded to impact the participants. The following sections set forth to understand how.

Both with regards to expected consumption and actual consumption, there is a statistically significant difference between the genders. The women expect to eat more and they also eat more than the men, even though they tend to overestimate their actual consumption to a higher extent than the men. Reasons for why this may be are discussed next. One proposition is that the primary outcome of the nudge is a raise of awareness in terms of actual consumption of fruit and vegetables. It can be argued that the men estimate their consumption quite accurately, whereas the women tend to either overestimate or be too optimistic about their consumption and must realize that the actual consumption does not

meet their expectations. Reasons for this difference might be ascribed to the social desirability bias, which proposes that certain people have a tendency to present themselves in a socially desirable way. We proposed already as a hypothesis that certain people will tend to state their consumption higher than it actually is in order to seem healthy, which can be ascribed to the social desirability bias. We argue that some of the women tend to state their consumption higher than it actually is, only to come to the conclusion that they are not able to live up to their expectations, hence the lower actual consumption.

Multiple explanations for why this may apply exist. In the following, we present a range of potential reasons but they are not discussed further, as this is not the scope of this thesis and the potential reasons are not empirically founded. Reasons for the differences between the genders might be caused by norms in society and a general expectation for women to be healthy. However, one might argue that such expectations for health and wellness do also apply for men today. Further, women might be stronger affected by the group especially within the sensitive subject of healthy eating, or they might make stronger demands on themselves and hence tend overestimate their consumption.

This tendency for women to overestimate is further supported by results of the control group. When looking into the numbers for actual consumption of the control group, we find that the women of the control group claim to have eaten on average 3.63. This is close to ¾ plate of fruit and vegetables and is larger than any of the other averages that we observe. The men of the control group claim to have eaten 2.78, which is very close to the average actual consumption for the men of the experimental groups (2.77). One could claim that the probability for women of the control group to have overestimated their consumption is quite large, as their actual consumption of 3.63 is far from the actual consumption of 3.26 of the women in the experimental groups. As mentioned in the methodology chapter, a delimitation of the experiment is that it cannot be known if we by coincidence include only very healthy women, vegetarians or vegans in the control group. If this is the case, it could mean that the actual consumption of fruit and vegetables of the women in the control group is in fact larger. However, we suppose that this is not the case and therefore we argue that these results could support the claim that the nudge functioned as an eye-opener which succeeded to raise awareness on the actual level of healthy eating.

In line with these thoughts on the raise of awareness, it would be interesting to see what would happen if we ask the same group of people about their consumption expectations and actual consumption on the following day or the following week. One could suppose that the participants and especially the women would come closer to a realistic answer for EC and hence minimize the difference between EC and AC. Another suggestion is that the nudge could have an even larger effect in week 2 as the women have become very aware of their consumption and the fact that they are probably not as healthy as they consider themselves to be.

On the opposite, one might claim that the nudge did in fact not have any effect on the women and that we only succeeded to nudge the men. It would be interesting to further investigate whether there are any general differences in the way that men and women react to being nudged and whether it could be a general tendency that women do not react as much to being nudged as men. At the moment, there is to our knowledge no research that goes into this field.

Another proposition for why there is a significant difference between the genders is that the men might to a greater extent be impacted by their commitments. It could be argued that the men feel more committed to their stated expectations, and hence feel that they have to live up to them. This is based on the fact that the difference between EC and AC ($|EACR|$) is a little smaller for the men than for the women. However, this difference is not statistically significant and therefore we cannot trust it to be true.

As a final proposition for why the difference between the genders is observed, it is realized that because the women on average expect to eat more than half a plate of fruit and vegetables, it might be wrong to increase the consumption even further, as such consumption might result in malnutrition in terms of a balanced diet of proteins, carbs and fat. For men, it would be relatively easier to increase the consumption of fruit and vegetables, as the starting point of expected consumption is smaller. It is likely that it is easier for the men to increase their consumption from a ¼ plate to a ½ plate of fruit and vegetables than increasing ones consumption from a ½ plate to a ¾ plate. Further, one could raise the question of whether nudging the participants to eat more than ½ or ¾ plate of fruit and vegetables would even count as an act of paternalism, i.e. is it really to make the

pants ‘better off’ to nudge them into eating a diet consisting of ¾ plate of fruit and vegetables or would this result in malnutrition. From this, it could be argued that the successful outcome of the nudge will differ from participant to participant, depending on the level of EC. In general, it could be argued that a successful outcome of nudging the men will be a small increase in consumption, whereas for the women a successful outcome will rather be consumption of the exact amount as stated. Of course, this is not true for all the women and all the men, as the expected consumption differs between the individual women and men but it is based on the overall result of the experiment.

Finally, we attend to the hypothesis that we developed prior to the experiment, which proposes that: *“women will be more affected by the nudge than men, i.e. the increase in in consumption will be greater for women than for men*.” As the results show that there is in fact an overall decrease in consumption for the women and a small increase for the men, we reject the hypothesis. From this, we should be able to accept the null-hypothesis saying that *“women will be less affected by the nudge than men, i.e. the increase in consumption will be smaller for women than for men”*. However, we propose that the proposed null-hypothesis might in fact not either be true. In this, we challenge our own first conviction that an increase in consumption equals a successful nudge and instead propose that the success is to a greater extent dependent on the level of expected consumption. This is again with a view to the fact that for a nudge to be successful according to its definition, it should be an act of paternalism that makes the participant ‘better off’. If looking to our considerations about what counts as ‘better off’ prior to this experiment, we found that ‘better off’ in terms of a healthy amount of fruit and vegetables would be around the 2/5 plate as recommended by the Danish Veterinary and Food Administration. In the experiment, we find that the most frequently observed value is ½ plate of fruit and vegetables, which indicates that the half plate is in fact closer to ‘better off’. With this knowledge, one could assume that nudging the participants to eat more than ½ a plate of fruit and vegetables would in fact not count as an act of paternalism, as such behavior would not make the participants ‘better off’. From this knowledge, one would be able to develop a new experiment for which the hypotheses could be adjusted accordingly and instead seek to nudge the participants to eat exactly ½ a plate of fruit and vegetables.

The above are all propositions that we believe can be the causal explanation for the statistically significant difference between the men and the women of the experiment.

Groups

In this next section, we look into the observable differences between the two experimental groups of the experiment, the 9 o'clock group and the 12 o'clock group, which initially had the primary focus of this research. Even though the analysis showed that the immediate differences between the two groups are not statistically significant, we still find that there are elements worth discussing, also considering the fact that the differences are quite different from what we expected to find through the experiment.

As we initiated the experiment, we worked from the hypothesis stating that the 9 o'clock group would have a larger tendency to overestimate their consumption, as they were asked to state their consumption several hours before lunch. We based this proposition on the idea that the participants of the 9 o'clock group would be full from their recent breakfast at the time that they were asked. Furthermore, the participants of the 9 o'clock group would have many hours ahead of them at work before getting to lunch and they would not to the same extent think of the consequences connected to the choices that they were asked to make at 9 o'clock. However, the overall result of the EACR for the 9 o'clock group shows that the participants in general underestimate consumption, whereas the 12 o'clock group tend to overestimate. This is opposite from what we expected to find.

On the other hand, when looking to the numerical value of EACR, it is found that the differences between EC and AC is smallest for the 12 o'clock group, which indicates that the 12 o'clock group is better at estimating their actual consumption compared to the 9 o'clock group. This might be a natural cause of the different times of the day that the participants are asked, which is in line with our thoughts about the differences between making an estimate around 9 o'clock versus around 12 o'clock. Further, it could be suggested that the smaller difference between EC and AC for the 12 o'clock group is a result of the availability heuristic. We propose that the participants might be biased in their decision-making because the statement of consumption is clearer in mind and hence easier to retrieve in the choice situation. This could then indicate that the nudge was in fact successful and that the availability heuristic proved to function as a means to nudge. However, as these re-

sults are not statistically significant we cannot trust that the observations did not just happen by a coincidence. Therefore, further research should go into a setup of this kind to be able to determine if and how the availability heuristic can best be applied and to broaden the understanding of how nudging at different points in time can best impact one's decision-making.

Another interesting aspect of the difference between the two groups is that the 12 o'clock group has lower expectations for consumption than the 9 o'clock group. Generally, the actual consumption of the 12 o'clock group is also lower than the actual consumption of the 9 o'clock group, despite that fact that the ｜EACR｜ shows that the misestimation is larger for the 9 o'clock group. We find it interesting that the participants of the 9 o'clock group seems to have higher expectations for consumption and that they do in fact also have a higher actual consumption than the 12 o'clock group. In the next section, we look into why such difference is observed, however, acknowledging the fact that the difference is not statistically significant.

First, it is proposed that the result could be an outcome of risk aversion. In general, making decisions involve a certain degree of risk. Choosing one thing naturally excludes other options. The degree of risk depends on the consequences of the decision and this impacts the risk aversion of the decision-maker. We propose that the 12 o'clock group might be more risk averse than the 9 o'clock group, and that this is the cause of the participants of the 12 o'clock group expecting to eat a smaller amount of fruit and vegetables than the 9 o'clock group. For the 12 o'clock group, the choice that they make will have an immediate impact, as the participants will have to make their actual choice for consumption of fruit and vegetables straight after they have stated their expectations for consumption. This might make the decision-making situation seem more risky to the participants compared to the decision that the 9 o'clock group has to make. For the 9 o'clock group, the consequences of the choice are probably not as clear, since the actual choice of consumption lies further ahead in time. A possible reaction to such feeling of risk might result in that the 12 o'clock group lowers their expectations to ensure that they can live up to their own statement. This might be a plausible reason for why the expected consumption is lower for the 12 o'clock group. Additionally, one could suggest that making a choice about consumption at 12 o'clock, when hunger, tiredness, and time pressure has kicked in relies to a

higher extent on the impulsive brain system in system 1. Therefore, the answer becomes an immediate choice. On the other hand, the choices made by the 9 o'clock group might be more well-considered and driven by system 2-thinking, which is the goal-directed brain system, as neither biological needs, riskiness and consequences or environmental cues are at play. At least not to the same extent as for the 12 o'clock group. Therefore, the 9 o'clock group will have more cognitive capacities to make a well-considered choice.

If this is the case, one could expect that we would observe a larger increase in actual consumption for the 12 o'clock group, as the participants might state a lower expectation to keep the risk at a minimum, and therefore they would easily be able to consume a larger amount than expected. This is however not the case, as the results show that in general terms the 12 o'clock group tend to overestimate their consumption. Additionally, the 9 o'clock group even underestimates their consumption. If proposed that the 9 o'clock group is less risk averse, one could suppose that the participants' statement of expected consumption would be rather optimistic but this seems not to be the case. Though, what is found most interesting is that even though the misestimation between EC and AC is larger for the 9 o'clock group, the 9 o'clock group did in general have a higher consumption of fruit and vegetables than the 12 o'clock group. If taking the design of the experiment into consideration, this means that the results of consumption for the 9 o'clock group is considered a better results in terms of eating a larger amount of fruit and vegetables, as the participants of the 9 o'clock group actually succeeded to live up to their expectations compared to the 12 o'clock group. One could then suggest that the nudge proved most successful on the 9 o'clock group, however still being aware that the results are not statistically significant. One proposition might be that if the participants of the 9 o'clock does in fact make a well-considered choice around 9 o'clock where they have a larger amount of cognitive capacity to make a rational choice, it could naturally result in that the expectations are closer to 'reality' and hence that the actual consumption proves to live up to expectations.

The next section develops further why this difference between the two groups is observed with a focus on the different heuristics and biases that come into play. Though, we are aware that the difference between the two groups is not statistically significant, we still find

it interesting to further investigate the immediate differences, as this is the primary scope of our research.

Isolation of the Heuristics and Biases

In this section, we take a deeper look into each of the heuristics and biases that we have come across in this thesis in trying to isolate if and how they have each helped to impact the decision-making of the participants of the two experimental groups. We start out by looking into the availability heuristic, which includes biases of retrievability and by imaginability. Hereafter, we move onto a discussion about commitment and the effects of committing oneself to an event. Finally, we look into the biases of social desirability, optimism and anchoring, and seek to understand how these have a general impact on the other events and biases that we observe.

Availability heuristic

Starting out with a look at bias by imaginability and the ease with which one can image an event, we propose that both groups are under the same conditions in terms of imaginability. First of all, the event that we ask the participants to predict is a well-known event that they meet every day, namely the act of going to lunch and choosing an amount of fruit and vegetables from the buffet. With the help of the five pictures, we make the decision-making situation even easier to imagine. Therefore, this should not be the cause of the difference between the two groups. However, one could suggest that the participants of the two different groups find themselves in two different microenvironments, and that this is the cause of the different answers for expected consumption.

The element of the microenvironment was added to the nudging definition by Hollands et al.. They suggest that: "*interventions (that) are implemented within the same microenvironment as that in which the target behavior is performed".* The microenvironment within which we operate is the canteen. This is the spatial microenvironment and this spatial environment is the same for both experimental groups. We suggest that a temporal element could be added to the microenvironment. In this thinking, we suggest that even though the physical environment is the same for both groups, the participants might find themselves in two different temporal environments based on the different times of the day. Therefore,

they approach the choice situation from two different mindsets. For example, when we ask the participants of the 9 o'clock group to state their expectations for consumption at lunch, they are on their way to work, probably not thinking about lunch, and hence, they are in a certain temporal environment. The environmental cues of this temporal environment, such as finding one's access card, grabbing a cup of coffee and going to the elevator, trigger the habitual actions involved in going to work. For the 12 o'clock group, the temporal environment that they find themselves in is fully associated with the problem area that we attend to, as they are on their way to lunch in the very moment that we ask them to state their expectations for their consumption of fruit and vegetables. Therefore, one could suggest that the 12 o'clock group has better prerequisites for making the decision for expected consumption, as both the spatial and temporal environment trigger the habitual actions that they make during the lunch break. In this, one could suppose that the 12 o'clock group would be more biased by imaginability as both the temporal, the spatial environment, and the pictures of the plates help them propose a valid estimate for their consumption. This could also help to explain why the misestimation is smaller for the 12 o'clock group than for the 9 o'clock group.

Further, this proposition could be backed by the ideas about bounded rationality, suggesting that one makes decisions based on the information that is available to you during the moment of the decision-making. Hence, the participants of the 9 o'clock group make the best possible estimate for consumption around 9 o'clock though being aware that much can happen between 9 o'clock and the lunch break, which could influence the actual consumption. We argue that the decisions made by the participants of both the 9 o'clock group and the 12 o'clock group are rational decisions, even though misestimates are observed. This is based on the knowledge of bounded rationality, as proposed by Simon. In this, we again find it interesting that the 9 o'clock group even underestimates consumption. This is another argument for why the decisions made at 9 o'clock can be evenly rational and as proposed, probably more well-considered and deliberate compared to the 12 o'clock group. Hence, the participants of the 9 o'clock group are not considered to be neither too optimistic nor irrational. This leads us onto a discussion of the commitment heuristic and how this might be the reason for why the participants of the 9 o'clock in general eats more than the 12 o'clock group.

Finally, as proposed earlier, the fact that the misestimation of the 12 o'clock group is in general smaller than the misestimation of the 9 o'clock might be ascribed to the availability heuristic, as proposed prior to the experiment. However, this difference between the two groups is not statistically significant, and therefore we cannot determine whether this observable difference is caused by chance.

Commitment heuristic

When looking to the element of commitment, we propose that the participants feel committed to their statement of expected consumption and that this influences the actual consumption. We propose that the higher consumption of fruit and vegetables of the 9 o'clock group might be caused by the fact that the 9 o'clock group also committed themselves to eat a larger amount of fruit and vegetables, which they might feel that they have to live up to. If this is the case, one should seek to determine when people are most inclined to set the highest expectations for themselves and hence state the largest expectations for the amount of fruit and vegetables, which would then give the best possible outcome. In this way, we can ensure that the participants eat a large amount of fruit and vegetables, as they do not want to fail to live up to their own expectations and goals.

From our results, it seems to appear that around 9 o'clock the participants are more inclined to set high expectations for themselves, compared to right before lunch. In this case, it would be better to insert the intervention in the morning around 9 o'clock than around 12 o'clock, in order to influence the participants in the best way possible. However, as these observations are not statistically significant, we cannot say with certainty that around 9 o'clock is better than around 12 o'clock. Though, we suggest that nudging the participants around 12 o'clock indicates to have impacted the participants negatively, as the participants of the 12 o'clock group eat less fruit and vegetables than the control group. This is supposedly a mere result of the nudge, as the participants of the 12 o'clock group had the same options available as both the participants of the 9 o'clock group and the control group. This means that the participants of the 12 o'clock group made their decision under the same conditions as the other two groups. In this thinking, we suggest that making the participants commit themselves to a certain goal under risky circumstances might prove to be rather un-paternalistic, as the participants might do themselves a bad

turn. As the participants supposedly lower their expectations because of the stronger feelings of risk, they might still feel inclined to keep their promise of consumption, despite the fact that the stated expectations for consumption are rather low. This results in the lower actual consumption, which in our case is understood as “unhealthy”. Therefore, we propose that when using the commitment heuristic to nudge people to better behavior, one should consider the fact that commitment might work differently under risky versus riskless circumstances. We suggest that further research should go into this to broaden the understanding of whether this could prove to be statistically significant in a similar experimental setup and whether the same logic applies in other circumstances.

Anchoring, optimism bias and social desirability bias

If looking to the final three biases that we have worked with through this thesis, we suggest that anchoring, optimism and social desirability can influence the strength of each other and hence have an indirect influence on EC and AC. We propose that the different biases come into function at different times of the decision-making process. For example, an anchor could be the starting point for the whole decision-making, which might be impacted by the effects of the optimism bias in terms of e.g. the participants’ estimate for EC, whereas the social desirability bias might influence the statement of actual consumption in terms of how accurately and honestly one responds. This indicates that the different biases are in fact interdependent. Therefore, we propose that it will be difficult to isolate the effects of the single biases, as they seem to impact each other across the whole decision-making process.

Evidently, we cannot isolate the effects caused by the biases of the availability heuristic in our experiment, as we have not been able to determine whether the immediate differences across the different variables are caused by commitment, anchoring, retrievability or any of the other biases that we have found to have an effect. Instead, we suggest that the different biases function interdependently and that one bias impacts other biases and might increase or decrease the effect of the other biases. In this, we mention again that the biases that we give attention to might not even be the full list of biases that have an influence on the decision-making situation. For example, we discovered yet another bias, the bias of

group mentality, as we conducted the experiment, which we did not take into account in the experimental design.

Can you Nudge a Habit?

In this section, we look into the element of habitual actions. We discuss the fact that the behavior that we set forth to nudge is found to be a habitual action, and we consider the consequences that this might entail. As mentioned, we propose that the consumption at lunch is driven by habits, which could influence the effect of the nudge. Therefore, we discuss whether nudging has a different impact on habitual versus impulsive actions, and whether it is more difficult to nudge in a stable environment versus an unknown environment.

First, looking to the question of whether nudging has a different impact on habitual actions versus impulsive actions, we attend to the two different core brain systems within system 1 as proposed by Vlaev and Dolan. We question whether it is even possible to nudge habitual actions or whether it is most appropriate to nudge only impulsive actions. As proposed by Vlaev and Dolan, one can nudge both habitual actions and impulsive actions but some heuristics suit each of the two actions more appropriately. First looking at habitual actions and the division of mental habits and motor habits, we propose that the actions that we attend to, i.e. the act of choosing an amount of fruit and vegetables at the canteen buffet, is a motor habit. This is in the sense that the act is learned through repeated practice in the same circumstances with the same states and stimuli being present.

Further, as already mentioned in this chapter, several findings point to the fact that the act is an act of habit. Such habits are usually the result of a desired goal, and with this in mind, Vlaev and Dolan propose that the best way to influence such a habitual action is for example with the use of the mental heuristic of commitment. Commitments can work to change or break habits, as they can oftentimes help people commit themselves to something different from the habit. Such a commitment might be necessary for people to break with an undesired or unhealthy action, which has become an inveterate habit after continuous repetition over a long period of time in the same circumstances. However, such thinking implies that the habit that you seek to influence is a bad habit, and that the Choice Ar-

chitect seeks to change it by having the participant commit himself or herself to a better act. From this, we raise the question of whether this habit that our experiment centers around is in fact one that we should seek to change. It is evident that for some of the participants, we can easily improve and influence their habits of consumption of fruit and vegetables, whereas for others, their habits are already 'good habits' and they should probably not be changed. This will be further discussed at a later stage, when discussing the act of paternalism. From this, it is evident that we can influence habitual actions, but with such actions follow another set of considerations that should be taken into account when developing the nudging intervention.

Again, as already proposed, the acts of commitment by the 12 o'clock group might be more of an impulsive action, as basic needs such as hunger and tiredness have kicked in. Therefore, we propose that the cognitive abilities are very limited and that this trigger an impulsive action. As mentioned, this might be the cause of the lower expectations of the 12 o'clock group. In this, we proposed that the nudge proved to have a negative effect on this group. This could indicate that nudging by commitment is not appropriate for matters of impulsive actions or under risky circumstances.

We argue that you can nudge both habitual and impulsive action but we support Vlaev and Dolan's claim that certain heuristics are more suitable for certain situations and actions. One could propose that in impulsive decision-making situations, one could benefit from applying very simple heuristics and alterations of the Choice Architecture to influence the decision-making. An example could be the classic canteen-nudge where small changes in the display and arrangement of food can change the consumption.

Additionally, we suggest that there will inevitably be differences between being met with interventions in a well-known and stable environment, where one acts on autopilot compared to meeting interventions in an unknown environment. For example, one might suggest that interventions in a stable environment might have a large impact in the beginning but at some point, the intervention will become a natural part of the environment. It would be interesting to investigate what this will mean for the effect of the intervention. On the other hand, in an unknown environment, the nudge might prove to have a longer lasting

effect, as it will not to the same extent fade into the environment. However, it could be proposed that the intervention might not be as disruptive as an intervention in a stable environment might be.

Furthermore, one might propose that the individual heuristics and biases work differently given the type of action that one seeks to nudge and the environment within which one intends to insert the alteration. For example, one could propose that commitments work differently within different subjects areas, amongst different crowds, and in different microenvironments. Within some environments, one might be keener to keep one's promise or within a certain subject area, which is considered less important, one might feel less committed to one's goal. In the same sense, the optimism bias and the social desirability bias might as well increase or decrease the power of other biases differently, depending on the subject matter, the environment, and the crowd. This is something to consider when developing the nudge. We propose that the nudges should be adapted accordingly whether the decision-making is habitual or impulsive and whether the environment is well-known or unknown. Further investigation should go into this field to be able to determine the potential differences in nudge effects within habitual versus impulsive decision-making processes.

A Paternalistic Nudge?

As already mentioned, we acknowledge the fact that increasing the consumption from ½ a plate to ¾ plate of fruit and vegetables might not actually count as being 'better off', and if so, the act cannot be considered paternalistic. In this sense, we already raised the concerns for whether nudging the participants around 12 o'clock proved to negatively impact the decision-making and hence the healthy consumption. Therefore, we question whether our act of inserting the intervention into the choice situation does in fact follow the definition of a paternalistic act. At least, we propose that the act should be altered to e.g. only be inserted at 9 o'clock to get the best possible results, i.e. the healthier results. Then, the nudge would count as a paternalistic nudge.

From this, it is evident that even after mapping out the different variables and taking a wide range of heuristics and biases into consideration, we were still not able to isolate the heu-

ristics and biases and the effects that follow. Despite the fact that we set forth to determine the effects of the different heuristics and biases, we were not able to isolate the causes of the different behaviors. We acknowledge that as the immediate differences in means are mostly not statistically significant, the differences in behavior might in fact be due to chance and that this might as well be the cause for why the heuristics cannot be isolated. However, we still propose that the immediate differences might be an outcome of the nudge, and hence we should to some extent be able to determine how and when the heuristics proved to have an impact. Further, for the difference between the genders, we are still not able to isolate the different heuristics and determine exactly what the causes for the difference are. One thing we can say is that the difference is not caused by the availability heuristic, which had the focus of the experiment. Finally, we propose that the fact that the effects of the heuristics and biases cannot be isolated is supported by the claim that the different heuristics and biases might in fact be interdependent.

If this is the case, it leads to a range of managerial implications in the form of obstacles for the application of nudging in real-life situations, e.g. as a means to impact decision-making during political campaigns or in organizational campaigns targeting both consumers and employees. First of all, nudging is obviously a very complex area, as it centers on psychological elements and human cognition. Understanding how this works is a rather demanding task, but to our conviction a necessity for making the intervention work. There might be risks of skipping this part in real-life situations and instead design nudges from seemingly clever ideas, which in the end proves not to function, as the theoretical basis is non-existing. More researchers seem to be of the conviction that any intervention can prove to be a nudge if it succeeds to improve the decision-making of people. This has resulted in that nudging can be anything. Convenient as this may be, it may also result in quite some risks for the effects of nudging. One could question if such clever ideas can live up to the element of paternalism and hence ensure that the nudge makes the consumers ‘better off’. Additionally, even after giving great attention to different heuristics that may apply in the choice situation at hand, and after taking into consideration the paternalistic element of the act that one sets forth to make, one might still not be able to understand the full ramifications of the intervention. Even if the nudge proves successful, one might not be able to determine exactly what the successful element is or what makes this exact setup successful,

which makes it rather difficult to apply the learning in future cases. This poses a general challenge for the efficiency of nudging as a marketing and communication tool. It is evident that nudging should be applied with caution and under well-considered terms if it shall live up to the element of paternalism.

Nudging is not be for the Masses

Following the above discussion, we attend to our concerns about how-to guides and nudging classes and the question of how they might negatively impact the power of nudging. If one cannot determine the exact effect of the nudge and herein the return-of-investment, we propose that nudging might not be for the common marketing and communication department. One could also question whether the paternalistic element of nudging can be fulfilled if applied uncritically in marketing departments and HR-departments to improve behaviors in organizations. We argue that nudging could in many instances be wrongly applied to impact behavior to the benefit of the firm and probably not the benefit of the consumers/employees. In such case, nudging is evidently not applied according to its definition. We propose that such faulty use of nudging might damage the general attitude towards nudging as an efficient communication tool. Furthermore, one could question whether nudging should at all be applied to all kinds of issues and communication problems, which seems to be the conviction within many marketing and communication departments today, or whether certain issues are not to be solved by the use of nudging.

In line with these concerns is as well the ethical aspect, which should always be considered when deliberately taking on the role as a Choice Architect. We already touched upon issues of the acts of paternalism within organizations and whether such acts actually have the ‘better off’-perspective in mind. When taking the result of our experiment into mind, we propose that we might in fact have impact the participants of the 12 o’clock group negatively, even when having the ‘better off’-perspective in mind. Therefore, we argue that similar events could easily happen in real-life situation. This proposes that making people ‘worse off’ under the intention of nudging might easily happen, and in such sense, nudging is definitely not ethical. On the other hand, we argue that if nudging is applied correctly, under the right circumstances and with the necessary knowledge, nudging can indeed

help consumers become ‘better off’. In such a situation where neither the libertarian nor the paternalistic element is violated, nudging can in fact be ethical.

Sub-conclusion

From this chapter, it is evident that the outcome of the experiment did not enable us to isolate the individual heuristics and biases and the effects that followed. With some certainty, we can point to several observations that indicate how the heuristics and biases enter into a dynamic web of influences. We have proposed a range of plausible reasons for the observations that we made through the experiment and the results that was found in the data analysis. These are all arguments bound by the heuristics and biases that we have attended to throughout the thesis, and they support the claim that the individual heuristics cannot be isolated, as they appear to be interdependent. The primary finding is that the genders react differently to being nudged. We propose a range of reasons for why this may be, e.g. based on the social desirability bias and the optimism bias but from the scope of this experiment we cannot say with certainty what causes the differences. Second, we propose that the immediate difference between the 9 o’clock group and the 12 o’clock group indicates that the two groups do as well react differently to the nudge. Even though this difference is not statistically significant, we look into why the immediate difference appears, as this has the primary concern of the thesis. Finally, from the findings we continued into discussions about habitual versus impulsive actions, as well as the question about paternalism, to broaden the understanding of our specific nudge and the consequences and challenges that this entails.

FUTURE RESEARCH

In this chapter, we bring together all the directions for future research that emerged from this research. These are all elements that can lay the ground for future research within the topic of nudging and especially nudging by the availability heuristic. First, we look into how one could give further attention to matters of nudging by availability by conducting an experiment that takes all the observations that we made during this research into consideration.

In general, we propose that future research should go into a similar setup, as the experimental setup that we developed, but running over a longer time-period. The purpose of this would be to determine whether the results of expected consumption, actual consumption and the difference between the two change over a longer time-period. In this one could e.g. investigate whether the optimism bias diminishes, if the commitment heuristic weakens etc.. We find that it would be interesting to observe the same set of participants over a period of e.g. a couple of weeks and look into how their responses evolve over time. One could suppose that the tendencies to overestimate/underestimate would fade over time and that the participants become more aware of their actual consumption. Investigation into this would strengthen the results across days, groups, and genders and create a stronger basis for concluding on the effect of the nudge.

Second, we propose that bringing in another topic for the experiment could enhance the understanding of the dynamics at play. One way to develop the understanding of how the different heuristics and biases work to influence behavior would be to apply the same experimental setup on a different topic than fruit and vegetables. One could suggest that if bringing in a topic such as beer or cake, where one could suppose that ‘better off’ is a smaller amount, the dynamics between the different heuristics might prove to be different. For example, one could suggest that the impact that this makes on different groups or at different times are different from these that we observed in the experiment of this thesis. Further, one could bring in a similar topic, such as a varied diet, or the consumption of whole grain or fish, to see if some of the same tendencies as we observe can be found. This is especially interesting with regards to the differences between the genders, which we observe in our experiment. If bringing up another topic, one could investigate whether

similar differences between the genders apply. For example, one could seek to broaden the understanding of whether women versus men tend to overestimate or underestimate in certain situations or within certain areas, with the purpose of determining exactly why this difference between the genders is observed. Further, it would be interesting to broaden the understanding of whether men and women react differently to being nudged.

Finally, we propose that further research should go into the effects of the availability heuristic as a means to nudge people in decision-making situations. We find that knowledge about when, i.e. what point in time, a nudge must happen for it to have the best possible impact on behavior is vital for the application of nudging in a marketing and communication context. Within our experiment, we propose that 12 o'clock was not the most appropriate time to nudge the participants, as we observed smaller expectations for consumption as well as smaller actual consumption within this group, which is considered the 'wrong' response. Such observation is a key finding for the use of nudging as an effective marketing and communication tool. Further research into this field should seek to determine exactly when a nudge must happen for the instance or memory to be available for it to have a decisive impact on behavior. We propose that the exact moment might vary from situation to situation and from topic to topic, and future research should as well go into this.

CONCLUSION

With this thesis, we set forth to broaden the understanding of nudging decision-making situations by heuristics and biases and we sought to investigate whether an individual heuristic and the bias that follows can be isolated to appropriately measure the effects of the nudge. With this, we sought to obtain a better understanding of how nudging can be used as a strategic marketing and communications tool, and we challenged the idea that nudging is a tool that should be uncritically used by the masses. We posed the research question: *"How can the availability heuristic be used as a nudge to impact decision-making, and how can one isolate and measure the effects of a specific heuristic? Further, what are the consequences of the uncritical and general use of nudging in mass society?"*. From the research question, we found it most appropriate to conduct an experiment of our own, to be able to distinguish the dynamics and effects of a real-life nudge. The experiment provided us with a deeper understanding of different elements of the decision-making situation and general challenges to bear in mind when seeking to nudge people's behavior.

The focus of the experiment was the availability heuristic, where we sought to investigate the effects of especially the retrievability bias at two different points in time. We acknowledged that other heuristics and biases come into play, e.g. commitment and optimism but with seven different hypotheses we sought to isolate the effects of the different heuristics through a thorough data analysis across the three variables, groups, days, and genders. From the experiment, it became clear that such an isolation that we set forth to identify was not possible. We found that the heuristics and biases impact each other interdependently and therefore their single impact on the decision-making situation cannot be isolated. In this, we propose that the nudge exploited several heuristics and biases and that they all contributed to the outcome of the experiment.

Already when designing the experiment, it became evident that a set of different heuristics and biases might have an effect on the behavior that we set forth to nudge. We identified five potential heuristics and biases that could come into play, and we acknowledged that other heuristics and biases that had not come to our attention might as well exist. However, we sought to come closer to an understanding of how the availability heuristic in itself had an impact on behavior and hence how it can be used as a means to nudge. This be-

comes relevant when nudging is applied in a managerial context. We proposed that if nudging is applied by marketing and communication departments, one will need to be able to determine the actual outcome of the nudge and in this the return-of-investment. This will require thorough knowledge of exactly how the nudge proved either successful or unsuccessful and hence an isolation of the heuristics and biases in action.

The main finding of the analysis, when seeing to the effects of the availability heuristic, is that there was no statistically significant difference between the 9 o'clock group and the 12 o'clock group. This was both in terms of misestimation (|EACR|) and increase/decrease (EACR) from expected consumption to actual consumption. Even though we observed some immediate differences, which indicate that the misestimation is smaller for the 12 o'clock group, who have most recently stated their expectations for consumptions, and hence where one could suppose that the retrievability bias has the strongest effect, we cannot say with certainty that the difference is not just due to chance. We did as well observe some other interesting tendencies across the two groups, e.g. that 12 o'clock might not be the most appropriate point in time to nudge for this specific case that we attend to, as some of the results suggest that the nudge did actually impact the participants of the 12 o'clock group negatively. Again, as these immediate differences are not statistically significant, further research should go into this field to obtain a better understanding of the dynamics at play.

The most interesting finding from the experiment proved to be the statistically significant difference between the genders, both within expected consumption and actual consumption. As the difference between the genders does not apply for EACR and |EACR|, we have not been able to determine exactly what causes the difference between the genders. However, we found that the difference between genders of the control group is not statistically significant, which is why we argue that the differences in EC and AC between genders are in fact an outcome of the nudge. Evidently, the genders responded differently to the nudge, which impacted their decision-making. We presented a range of potential reasons for why this applies but further research should as well go into this difference to be able to determine the primary cause of the result.

Finally, it is evident that the experiment sets forth to nudge the participants into being 'better off', i.e. being healthier by eating a large amount of fruit and vegetables. Therefore, we considered the nudge to be an act of paternalism. With considerations about what is found to be 'better off' in terms of fruit and vegetables, also taking into account the recommendation from the Danish Veterinary and Food Administration, we were of the conviction that we could make the participants commit themselves to an optimistic goal, and with this goal fresh in mind, we believed that the participants would eat a large amount of fruit and vegetables and to some extent even eat more fruit and vegetables than expected. Though, it became clear to us that the nudge might in fact have had a negative impact on the consumption of the 12 o'clock group based upon ideas about risk aversion and commitment. Therefore, despite the fact that we considered our acts as paternalistic, we might in fact have done the participants a bad turn. This implies that much investigation should undergo the design of the nudge when applied in real-life situations to ensure that the intentions to nudge does not accidentally result in a negatively impact on the behavior of the consumers.

From this knowledge, we propose that nudging, as the general masses think of it today, is rather misunderstood. To our conviction, one cannot learn to be a nudging expert from reading a 4-step guide to developing "nudges that work". Furthermore, marketing and communication departments should as well not throw themselves into an uncritical use of nudging in an attempt to solve various communication and marketing issues. We propose that with such uncritical use of nudging, a range of consequences might follow. For example, one might risk that the nudge proves not to be paternalistic, i.e. the change in behavior does not make the consumer 'better off'. Second, if the marketing and communication department fails to be critical about the use of nudging, or fails to invest the right amount time into research of the nudging design, the nudge might prove unsuccessful. This again distorts the understanding of nudging as an effective marketing and communication tool.

With this research, we provide the marketing and communication professionals with a better understanding of how and when nudging can and cannot be applied as an effective marketing and communication tool. In this, we direct the attention to both opportunities and challenges that marketing and communication professionals can meet in the application of

nudging. We legitimize taking a critical account to the general tendency to think of nudging as the answer to all kinds of communication and marketing issues.

REFERENCES

Bornstein, B. H. & Emler, C. (2000). Rationality in Medical Decision Making: A Review of the Literature on Doctors' Decision-Making Biases. *Journal of Evaluation in Clinical Practice, 7, 2, pp. 97-107*.

Business Dictionary. (2018a). *Framing*. Retrieved April 23th 2018 on: http://www.businessdictionary.com/definition/framing.html

Business Dictionary. (2018b). *Decision Making*. Retrieved March 2nd 2018 on: http://www.businessdictionary.com/definition/decision-making.html

Chartered Association of Business Schools. (2015). *The Purpose of of the Academic Journal Guide*. Retrieved January 8th 2018 on: https://charteredabs.org/academic-journal-guide-2018/

Cherry, K. (2017). *Understanding the Optimism Bias*. Retrieved March 26th 2018 on: https://www.verywellmind.com/what-is-the-optimism-bias-2795031

Creative Research Systems. (2012). *Sample Size Calculator*. Retrieved January 17th 2018 on: https://www.surveysystem.com/sscalc.htm

Crowne, D. P. & Marlowe, D. (1960). A New Scale of Social Desirability Independent of Psychopathology. *Journal of Consulting Psychology, Vol. 24, No. 4, pp. 349-354.*

Easterling, R. G. (2015). *Fundamentals of Statistical Experimental Design and Analysis*. New Jersey: John Wiley & Sons.

Encyclopædia Britannica. (2018a). *Libertarianism*. Retrieved February 8th 2018 om: https://www.britannica.com/topic/libertarianism-politics

Encyclopædia Britannica. (2018b). *Paternalism*. Retrieved February 8th 2018 on: https://www.britannica.com/topic/paternalism

Evans, J. St. B. T. (2008). Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition. *The Annual Review of Psychology, 59: 255-278.*

Fischer, R. J. (1993). Social Desirability Bias and the Validity of Indirect Questioning. *Journal of Consumer Research, Vol. 20, No. 2, pp. 303-3015.*

Fødevarestyrelsen. (2018). *Spis frugt og mange grøntsager*. Retrieved January 16th 2018:http://altomkost.dk/deofficielleanbefalingertilensundlivsstil/de-officielle-kostraad/spis-frugt-og-mange-groensager/

Gigerenzer, G. & Gaissmaier, W. (2011). Heuristic Decision Making. *The Annual Review of Psychology, 62: 451-482*.

Hansen, P. G. & Jespersen, A. M. (2012). *Skraldespanden set indefra*. Retrieved March 10th 2018 on: http://www.kommunikationsforum.dk/artikler/nudging-paa-gadeplan

Hansen, P. G. (2016). The Definition of Nudging and Libertarian Paternalism: Does the Hand Fit the Glove?. *The European Journal of Risk Regulation. Vol 1, pp. 1-20*.

Kahneman, D. (2011). *Thinking, Fast and Slow*. London: Allen Lane.

Kahneman, D. & Tversky, A. (1984). Choices, Values, and Frames. *American Psychologist, Vol. 39, No. 4, pp. 341-350.*

Kvale, S. (2013). *Doing Interviews*. London: SAGE Publications Ltd.

Münster, M. (2017). *Jytte fra Marketing er desværre gået for i dag*. København: Gyldendal Business.

Nygaard, C. (2013). *Samfundsvidenskabelige analysemetoder*. Frederiksberg: Samfundslitteratur.

Simon, H. A. (1990). Bounded Rationality. In J. Eatwell, M. Milgate, & P. Newman, *Utility and Probability* (pp. 15-18). London: Palgrave Macmillan.

Thaler, R. H. & Sunstein, C. R. (2009). Nudge: *Improving Decisions About Health, Wealth. And Happiness*. Penguin Books.

Thaler, R. H. & Sunstein, C. R. (2003a). Libertarian Paternalism. *American Economic Review, Vol. 93, No. 2, pp. 175-179*.

Thaler, R. H. & Sunstein, C. R. (2003b). Libertarian paternalism is not an oxymoron. *The University of Chicago Law Review, Vol. 70, No. 4, pp. 1159-1202*.

Tversky, A. & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science, New Series, Vol. 185, No. 4157, pp. 1124-1131.*

Vlaev, I. & Dolan. P. (2015). Action Change Theory: A Reinforcement Learning Perspective on Behavior Change. *Review of General Psychology, Vol. 19, No. 1, pp. 69-95.*

Vlaev, I., King, D., Dolan, P. & Darzi, A. (2016). The Theory and Practice of “Nudging”: Changing Health Behaviors. *Public Administration Review, Vol. 76, Iss. 4, pp. 550-561.*

Weingarten; et al. (2016). From primed concepts to action: A meta-analysis of the behavioral effects of incidentally presented words. *Psychological Bulletin Vol. 142, pp. 472–497*.