



## **Judging about health nudging**

*Do objections raised against the general use of nudges  
also hold as strongly for health-related nudges?*

Charissa Rentier

Student number 329481

Bachelor thesis

Faculty of Philosophy

Erasmus University Rotterdam

Thesis supervisor: dr. Constanze Binder

Thesis advisor: dr. Conrad Heilmann

*Abstract:* In times of unhealthy food being plentifully available and exercise during our daily lives existing only in walking to and from the car, obesity and other physical, but also psychological, diseases reign. Living unhealthily often leads to long-term results many people did not have in mind, such as obesity or other diseases related to unhealthy life styles. It is thought the government, employers and other choice architects could help people make better decisions by nudging them towards those options that are more consistent with their own long-term goals. However, objections have been raised from an ethical point of view against the use of nudges. This thesis sets forward three important concerns, namely regarding autonomy, a lack of information on preferences and the heterogeneity of these preferences, and regarding truth and transparency. The objective is to investigate whether the objections against nudges are still as strong when it comes to health-related nudges. In particular, it is argued that the autonomy problem is less stringent, because of the greater ease of habit-formation in health, and autonomy may even be enhanced though nudges if they improve health. Also, because of the nature of health and the relative rather than absolute goal of the health nudges at hand, informational and heterogeneity problems regarding preferences are less severe, and the intrusion is weakened. The general objections against nudge do not disappear completely, but are less severe and not as worrisome in the health-related context as opposed to other applications of nudge. Since not all objections are discussed, it is recommended that further research focus on whether other objections actually become more severe in the case of health-related nudges.

*Keywords:* nudge, libertarian paternalism, health, ethics, autonomy, choice architecture, preferences, habit formation, transparency.

## Table of contents

1. Introduction	3
2. Theoretic framework	5
2.1 Nudges	5
2.1.1 Types of nudges	7
2.2.2 Areas of application	9
2.2 Health	11
2.2.1 Definition and determinants	11
2.2.2 Examples of health-related nudges	11
3. Permissibility of nudges in general	14
3.1 Autonomy	14
3.2 Information about preferences	15
3.3 Truth and transparency	16
4. Health-related nudges: what is it that makes health-related nudges a less objectionable type of nudges?	18
4.1 Relative rather than absolute goal of health-related nudges lessens the informational problem and intrusion	18
4.2 Habits are more easily created in the health context, leading to more intrinsic motivation to stick to long-term goals	22
4.3 Rather than impairing it, health nudges may even save autonomy	25
5. Conclusion	28
6. Literature	31

# 1. Introduction

The average person frequently displays inconsistent decision-making behaviour: people make decisions in the short run that taken together do not add up to the outcome they desire in the long run. Pension savings are structurally too low, environmental concerns are not translated into environmentally friendly behaviour, etc. In the extremes, people could either be forced to act according to their goals, or they could be left to work things out on their own; this is exactly the distinction between paternalism and liberalism. Another option, lying in between, is using nudges. Nudges are ways to influence decision-making by gently pushing people towards the option that would make them better off, thereby trying to get people to act more in accordance with their long-term goals while not restricting the options available to them. An example of such a nudge in the case of environmental concerns is to install a device in houses that shows households' energy consumption compared to their neighbours'. A green light could be used if their usage is relatively low, and a red one if they perform badly compared to their neighbours (Thaler & Sunstein, 2008).

An obvious area of application of nudges seems to be health, which is the topic of this thesis. In times of unhealthy food being plentifully available and exercise during our daily lives existing only in walking to and from the car, obesity and other physical, but also psychological, diseases reign. Many people live busy lives and do not want to spend the time they have exercising and preparing healthy meals, which generally takes longer than going for some takeaway. Living unhealthily often leads to long-term results many people did not have in mind, such as obesity or other illnesses related to unhealthy life-styles. Nudges, gentle pushes in the right direction, could be implemented to make people exercise more often and/or eat more healthily, which makes it more likely they realise their goals.

Developed by Thaler and Sunstein (2008) for policy-making purposes, the idea of nudge has managed to find its way into various governmental institutions and their policies. In July 2010, David Cameron installed the Behavioural Insights Team – commonly known as the 'Nudge Unit' –, a group of academia, policy makers and marketers. Their goal is to find innovative ways of encouraging, enabling and supporting people to make better choices for themselves (Behavioural Insights Team, 2010). To accomplish this, they use insights of behavioural economics to influence British policy-making on various topics, such as food waste, health and energy use. In the United States, where Cass Sunstein himself was head of the White House Office of Information and Regulatory Affairs until 2012, the nudge was at the base of the health care reforms president Obama introduced, and Sunstein also put various laws into place based on this insight from behavioural economics (Bates, 2012).

The concept of nudges and also the implementation of various nudge-based policies gave rise to various types of research – next to the experimental settings in labs –, looking at the potential effects of the nudge in various areas and at which mechanisms are most effective. Some examples of research showing the effectiveness of nudges are Mitchell & Moore (2011) on lifetime individual retirement arrangements, and

Gillingham & Palmer (2013) on bridging the energy gap by using insights from economic theory and empirical analysis.

There are, however, some objections against nudging people towards certain options when it comes to (moral) permissibility. Previous studies have focused on discussing general problems, such as a piece by Luc Bovens (2009) in which he distinguishes three types of problems: people may not actually think about the change in behaviour they display due to nudges; the question of who is nudging; and the problem of transparency. Others, like Blumenthal-Barby and Burroughs (2012), have discussed the ethics of using paternalism and nudges in the health context specifically. Their main finding is that the degree of moral permissibility depends on how exactly the nudge takes form – so the type of the nudge – and on how it is applied.

The objective of this thesis is to investigate how the most important objections against nudges change when it comes to the application of nudges to health-related behaviour such as exercise and nutrition. It shall be argued that some of the objections raised against nudges are less severe in the case of health-related nudges. This objective is where the contribution of this thesis lies, since most of the literature focuses solely on nudges in general or nudges in health-related contexts, but does not combine and compare the two.

In order to investigate this change in the most important objections, the concept of nudge in its wider framework and the most commonly raised problems will be investigated through literature research. After this, the objections that are weakened in the case of health-related nudges are explored and an argument is made for this change in strength, using intuition, reasoning and theoretical argumentation.

Firstly, in the theoretic framework the concept and objective of nudge will be explained, together with the types of nudges and areas of application. Also, the definition and determinants of health that will be used in this thesis and also the possible applications of health-related nudges will be discussed.

In the main body of the thesis, the general moral permissibility of using nudges will be examined by looking at the most important objections raised against them. After this, it will be shown how the analysis of nudges changes when it comes to applying them to health-related behaviour.

Finally, the most important findings of the thesis will be summed up. Of course, not all problems related to the concept of nudge are weakened when applying them in a nutrition and exercise context, and therefore, the conclusion will briefly discuss the problems that remain. Furthermore, some limitations of the study and recommendations for further research are given.

## 2. Theoretic framework

This chapter contains the various theoretical concepts necessary for the analysis of health nudges. The concept of the nudge and libertarian paternalism are described. Furthermore, the forms these nudges can take and the various fields of application are briefly discussed. Next, health is defined and the determinants relevant for this thesis are explained. Lastly, examples of nudges in the health context are given.

### 2.1 Nudges

In discussing nudges, the libertarian paternalism from which it arose cannot remain unexplained. *Libertarian paternalism* is a term introduced by Thaler and Sunstein around 2003, combining libertarian and paternalistic intuitions in one view primarily used for policy-making purposes. Paternalism aims at protecting people from themselves by intervening in their lives ‘for their own good’ (Shafer-Landau, 2005). The paternalistic part of the concept shows itself in the aim to ‘nudge’ people towards a certain option in order to make them choose in a way they, looking back, also deem to have made better off than the option they would have chosen without the nudge. The libertarian part about the nudge is that the individual retains the freedom (relative to the purely paternalistic approach) to choose any of the alternatives. This means that even the least preferable alternative – that is: from a policy-making perspective and taking people’s long-term preferences into account – remains a positive option to choose at incurring very little costs in terms of energy or money. When it comes to nudges, autonomy and liberty are important: people should not be forced to choose the preferred option (Thaler & Sunstein, 2008).

According to Thaler and Sunstein this libertarian paternalistic approach is desirable because, contrary to the claims standard economic theories make about rationality, individuals do not seem to have stable and well-defined preferences, and their choices are influenced by the way the various options and the choice itself are presented to them. To make this less abstract, one could imagine the case of someone being on a diet. This person wants to lose weight or eat more healthily, but presented with unhealthy but very tasty food, he or she might give in, even though in advance he/she would not have expected to do so. This would mean that the libertarian approach, which assumes that people know what is best for them, that they are able to make these choices themselves and no intervention should take place to manipulate their decision meaning, is based on an erroneous assumption.

When it comes to choices and decision-making, choice architecture is always present: the context in which decisions are made. For example, while walking in the supermarket, the way the aisles are set up and the food is arranged on the shelves influence our shopping and even our choice behaviour. The person designing the choice structure is called the choice architect, and he or she always chooses a certain frame in which the choice is to be made, each aspect of this frame influencing individuals making decisions within it. This inevitable choice for a certain frame makes neutrality of the design impossible, because people making the decisions within such a structure will always be somewhat driven to a specific alternative, or to a certain mode of decision-making.

Though many definitions of nudge exist, the one that comes closest to what is meant in this thesis is one by Hausman & Welch. They define nudges as “ways of influencing choice without limiting the choice set or making alternatives appreciably more costly in terms of time, trouble, social sanctions, and so forth. They are called for because of flaws in individual decision-making, and they work by making use of those flaws” (2010, p. 126). What lacks in this definition is that manipulation cannot be called a nudge, so what should be added is that in addition to preserving all choice options, people should also not be forced to pick a specific option. Furthermore, nudges work best in the dark, as Bovens (2009) also says: people should be unaware of the effects they have, but at the same time have to be able to figure out what is happening (more on this topic will follow in section 3.3).

In a piece on the ethics of nudge, Luc Bovens (2009) states that nudges typically try to correct for agency problems, which are an example of the flaws in individual decision-making mentioned in the definition above. With this formulation, Bovens refers to the friction between a principal and an agent, where the principal is the planner and the agent is the doer. Evidently, sometimes the agent has other interests than the principal, which is the core of the agency problem. In the case of health nudges this could for example be if the principal, i.e. the self that has a long-term vision, wants to go to the gym twice a week, whereas the agent, i.e. the short-sighted self, wants to watch its favourite show and sit on the couch because it feels tired after a long day’s work.<sup>1</sup>

Bovens mentions six types of agency problems we face that nudges try to correct for (p. 212):

- (i) *ignorance*: when our lack of knowledge prevents us from reaching our goals;
- (ii) *inertia*: we have the knowledge but just put off the action;
- (iii) *akrasia*: we have the knowledge, we want to act on it, but when it comes to making the decision we lack the willpower to act in accordance with our long-term goals;
- (iv) *queasiness*: we just don’t want to think about certain things, because there is an emotional cost attached to it, for example in the case of organ donation after death;
- (v) *exception*: when we think we will be the exception to the statistical rule (Bovens gives the example of a sex change surgery to illustrate this point: if a person really wants it, he [or she] will want to go through with it, even if he knows that most of the people who undergo the surgery regret it afterwards, because he thinks he is not like the rest and is very sure about his decision);
- (vi) *social benefits*: when individual choice behaviour on an aggregate level does not lead to socially optimal results.

---

<sup>1</sup> For a brief overview of multiple-self models, see Wilkinson, N. (2008). *An introduction to behavioral economics*. Basingstoke: Palgrave Macmillan, pp. 242-246.

### 2.1.1 Types of nudges

In discussing the ethics of using nudges in health care, Blumenthal-Barby and Burroughs mention different forms nudges can take on and also describe which factors pose a risk of moral impermissibility (2012). Some of the options for applying the nudge are:

- ❖ Using *incentives*, often financial, to trigger the desired behaviour. These incentives can have many forms and can be positive, like a reward if you behave as wanted, or negative, like a punishment if the desired behaviour is not realised. Because these incentives motivate people to choose a certain option, but do not limit the options available, this form of influencing behaviour can be called a nudge. In a health context this type could be applied by making fast food relatively more expensive than healthy food, thereby motivating people to eat more healthily.
- ❖ Changing *defaults* (standard options) is another way of nudging, often mentioned by Thaler and Sunstein (2008) in the context of insurance and pension savings. This way of nudging appears to be effective because of the fact that people would rather outsource certain decisions and are therefore, as well as due to other psychological biases, inclined to choose the standard option if it seems sufficiently agreeable, thereby ignoring the other options. Again, options are not limited, but the most desirable option is made more easily accessible. Applying this to the health context, a default menu in restaurants could be changed to include less unhealthy food and more of the healthy type.
- ❖ Relying on *salience and affect*, in which images and representations play a major role. The idea is to influence people by novel, personally relevant, or vivid examples and explanations through images, narratives, etc. The goal is to evoke an emotional reaction, which afterwards shapes decisions and behaviour. To have an effect on behaviour, there should be a certain level personal relevance of the information shown. This type of nudge also leaves people free to choose, they are simply presented with more information or pictures that come to mind more easily when presented with a certain choice. When it comes to nutrition and exercise, an example would be images of healthy food or people exercising.
- ❖ *Norms and messenger* can also be used as a nudge, as it includes a message of what other people do – something to which individuals are highly susceptible, especially if they feel a connection to those people (Yun & Silk, 2011). Usually, a messenger is chosen because of certain characteristics, such as expertise or fame, making them more sympathetic and influential. In the health context, especially experts of food and/or exercise supposedly have a big influence, again not restricting options, but just guiding people to live more healthily.
- ❖ *Subconscious priming*, which is influencing thoughts and behaviour by giving subconscious cues, for example by making things more easily accessible, while at the same time leaving the less preferred

options still available. An option in the health context would be to make the stairs more easily accessible than the elevator, or to place the healthy food at eye height level while storing unhealthy snacks on very low or high shelves.

- ❖ *Commitments and ego preservation*, which responds to people's desire to stick to their commitments and to act in ways that benefit their self-image. In this form of nudge, self-esteem, but especially social status is at stake if the individual does not live up to their own long-term goals. The effect of this type of nudge is caused mostly by feelings, while at the same time not limiting options. In a health-related situation, this effect could for example be created by telling friends and family you are trying to lose weight. Then you would not want to be caught eating chocolates at birthdays while they are present, so you would more easily stick to behaviour that is in line with your long-term goal.
  
- ❖ Another option, not mentioned by Blumenthal-Barby and Burroughs but somewhat related to salience and affect and also found in different applications of nudge, is *disclosure*: informing people, mostly about their behaviour and its consequences, but also about the products they use and hardly ever think about. In informing people about the choice options, disclosure helps them update their knowledge and thereby forces them to make a new deliberation – different from the last time they made a similar decision. In addition to this informational point, making things more *concrete* rather than abstract would also be a good nudge: making explicit what the consequences of certain actions are, rather than using vague terms. In nutrition, for example, this could be done by providing people with more information about the nutritional values of products and about the effects things like too much salt and sugar could have on your body. People would then include this information in their consideration, thereby being more easily scared away from unhealthy products.

The six types mentioned by Blumenthal-Barby and Burroughs can be divided into two groups, which are *lower-order nudges* and *higher-order nudges* (Castelo, Reiner, & Felsen, 2012). This distinction is based on an analysis of human decision-making in which two systems are recognised: System 1, the intuitive system which operates automatically and the decision-maker has little control over, and System 2, the reflective system which deals with effortful mental activities such as deliberate choice and concentration (Kahneman, 2011). Lower-order nudges, such as those based on *salience and affect* or *subconscious priming*, intervene at the System 1 level, whereas higher-order nudges, such as *incentives* or *commitments*, have to do with deliberate choice by System 2. Nudges based on *defaults* and *ego-preservation* are somewhere in the middle.

This distinction is useful for evaluating the extent to which nudges infringe on the autonomy of decision makers. Lower-order nudges are assumed to have larger consequences for autonomy than higher-order nudges, in which people still engage in deliberate decision-making. In sections 3.1 and 4.3 the subject of autonomy will be discussed in further detail.



### 2.2.2 Areas of application

The different forms of nudges described in the previous section can be applied to many different settings in which people could use some help to choose ‘better’ options. In their book, Thaler and Sunstein (2008) mention various areas of application and how nudge are put to use in these settings. Here, in order to show that not just health-related behaviour is a viable target, some other areas of application and the reason for using the there are discussed below.

#### - *(Pension) savings*

When it comes to savings, and especially pension savings, standard economic theory (SET) makes some assumptions that do not seem to hold in real life. For example, with savings, SET assumes people will calculate their lifetime earnings, estimate what they will need when they are retired, and adjust their spending and saving pattern in the present accordingly. However, this assumes people are able to make these calculations (and also know when they are going to die), and that they have the willpower to stick to their plan, both of which are not very likely to hold for all people.

To make sure people save at least some money instead of none at all, because you often have to actively choose to save, nudges can be used. To achieve this, it is suggested to change the default option, which means people are automatically enrolled in a certain savings programme, rather than having ‘not saving’ as the default option and requiring people to actively subscribe to a savings programme.

#### - *Investments*

Like choosing a savings plan, investing is not something many people would do every day. Therefore, it is very difficult for them to decide between the various options, choose how much money to invest, and consider the risk and return all options entail. Even investors sometimes do not know what to do exactly, so they use rules of thumb.

To make investing a less tedious job, various nudges could be introduced to change the choice structure. An example is to change the default option in which portfolios are selected based on risk preference, age and other characteristics of the participant, so it will match their preferences better and makes the task less tedious. This can also be done by structuring the complex choice at hand, making more options available for very involved people, and fewer options for those who do not want to be. Another thing to work on is to explain to people how these abstract numbers translate into consequences later in their lives, so it becomes clear where adjustments can be made according to people’s wishes and preferences (disclosure and salience and affect are types of nudges that could be used here).

#### - *Credit markets*

In credit markets people particularly do not seem to act according to their own long-term goals, but also not in their own present interests. People borrow at high interest rates while they have money in their bank accounts on which they receive a lower percentage, do not look closely enough at the conditions and not far enough into the future to see the consequences. Examples are credit cards and student loans.

In order to prevent damage, people can be nudged to repay more of their credit card debts than they have to to prevent large amounts of money going to interest payments, for example by making them more aware of the difference this would make by presenting them with financial figures.

- *Organ donation*

Due to the relatively low number of people registered as donors, there is an enormous, continuously growing, waiting list for people who need an organ transplant, and many people die while waiting. Statistics from the United States show that the number of patients on the waiting list in the year 2006 was over 95,000, while the number of patients dying while waiting was over 6,300 (Abouna, 2008). Statistics over 2011 show an average of 18 people dying each day because of the shortage of donated organs (Donate Life New York, 2013). These figures indicate that there is a big problem, which could largely be solved if people who are willing, but haven't done so yet, would sign up as donors. However, few people like thinking about after-death organ donation, and of those who do think about it, some really object to it.

Since the default option usually is that people are not a donor unless they explicitly make known they want to. This requires effort, and since we usually do not think we will die anytime soon, we fail to make the effort to register as a donor. Those people who want to register, but put it off until later, could be nudged into registering by changing the default option: everybody would then be donor unless they actively make it known they do not want to be one. This would mean a shift from explicit consent to presumed consent; those who really object to being a donor will make the effort to unregister, while those who are indifferent or in favour of being a donor will not have to undertake action. However, it seems to be objectionable to assume people 'want' to be a donor. Therefore, a less objectionable nudge might be more feasible, such as mandated choice: forcing people to make a choice, so it would require the same amount of effort to either consent to or to reject becoming a donor.

- *Environment*

This area is a good candidate for nudges, because behaviour is not directly linked to consequences, and many of us will not live to experience them. There is therefore little incentive for us to take proper care of the environment, aside from internal motives such as genuinely caring about future generations and the effect your own behaviour has.

Since we are not all members of Greenpeace, this is a perfect occasion to use nudges in order to make people behave more environmentally conscious. This could be done by giving people more feedback and disclosing more information, for example by making comparisons of their energy usage and that of the neighbours. Another option is to make the energy savings more tangible in monetary terms.

Of these examples of areas in which nudges are applied, especially the ones about savings, organ donation, and the environment will be contrasted with the health application in the remainder of this thesis. The choice for these three subjects can be explained by the fact that governments generally pay more attention to them and because they are of a different nature, making comparisons with health nudges more interesting.

## 2.2 Health

### 2.2.1 Definition and determinants

Health has different definitions in various contexts, and should therefore be clarified before nudges concerning health are discussed. In investigating the topic at hand, it is important to choose a definition, and, more importantly, the determinants chosen as the focus of health nudges. Hoeymans, Melse and Schoemaker (2010) argue that health does not only have to do with statistical information, but also with the subjective experience of people within society. So, in addition to life expectancy, mortality and illness, it is also about the weight people attach to their health and their subjective experience of illness. Hoeymans, Melse and Schoemaker identify three determinants of health in this context: personal factors, behaviour and environment. The personal factors cannot be influenced easily, since they have to do with certain predispositions through genetic circumstances. However, behaviour and environment are two determinants that can.

In this thesis, the focus will be especially on individual behaviour, namely eating and exercise habits, since this is the determinant to which nudges can most easily be applied. Naturally, health is a word used for both physical and mental states, but here, the physical element in developing nudges will have a central position, and specifically the part influenced by exercise and nutrition. Mental states will also be influenced, but indirectly: various studies have shown that physical fitness leads to enhanced psychological health in terms of improved self-concept and self-esteem (Plante & Rodin, 1990), as well as better results in terms of self-efficacy and locus of control (Waller & Bates, 1992). In short, being healthier makes people feel better and more in control of their lives.

### 2.2.2 Examples of health-related nudges

Since health is socially relevant, it is an obvious area for applying nudges. In another bachelor thesis, various options for using nudges to change health behaviour related to exercise and nutrition were investigated (Rentier, 2012). Here, these will be discussed briefly in order to get an idea of which kinds of applications are possible in this context.

- ❖ *Increasing awareness and information* in order to become more aware of the factors that influence health and get more specific information about nutrition and exercise. This could include laws that force producers to mention nutritional values and other specific information on the food packaging. Also campaigns that promote exercise, for example at the municipal level, could help people to start working out more often. Another example is to provide information about a healthy life style. An important aspect is making this as specific as possible, for this is more likely to influence behaviour than a vague description of the consequences of exercising twice a week.

- ❖ *Changing the choice architecture and making the healthier option more easily available.* This could be achieved by putting healthy products on shelves at eye level in the super market, by making healthy food more accessible in the work place, designing buildings in a way that makes taking the stairs less cumbersome than taking the elevator, etc.
- ❖ *Insights from behavioural economics and psychology* could be used for example to increase the chance of people creating healthy habits, which could also be combined with economic incentives to make healthy food relatively cheaper than unhealthy food. Additionally, monetary rewards and punishments could be used to make people act more healthily. The rationale behind this example is that we could combine the knowledge about imperfections in decision-making with other psychological aspects of choice and economic incentives to change behaviour.
- ❖ Correcting the inability to adjust short-term behaviour to long-term goals by enforcing *commitment*, for example through contract-like arrangements, or by some form of social control. Even though contracts themselves may not seem like nudges, it is the option of the contract, added to the other options, mainly the option of no contract, which people can freely choose. Once they choose this option, make their goals explicit and agree on the consequences of not attaining that goal, people are more likely to commit to that goal than if they had not chosen the contract.

This could be done, for example, in the form of contracts in which people commit themselves to working out twice a week and defining consequences if they do not achieve this goal. Another example is the case in which people design a week menu at the start of the week, which they will have to stick to. Bodily influences like cravings for chocolate or other snacks will be less present if this schedule is designed on Monday morning after breakfast for example, than if they only decide seconds before actually eating. When Thursday afternoon comes and the craving starts, there will already be a schedule stating what and when you can eat, to which you can force yourself to stick.

A different example people can use to make themselves stick to certain goals if they know they will not do so on their own: stickK.com. This is a website developed in 2008 to promote a healthier lifestyle by enabling people to create so-called commitment contracts, which should make them reach their personal goals ([www.stickk.com](http://www.stickk.com)). On this site, people can enter their personal goals and voluntarily put a monetary amount at stake which they have to pay if they fail to meet their goal. They can also assign a referee to check their results, and it is possible for them to notify friends and family of their goals, so they do not only have to give account of their progression to themselves, but there is also some social control. An analysis of the 125,000 contracts that have been set up between 2009 and 2012 shows that appointing a referee makes the success percentage rise from 29 to 59 percent. However, putting money at stake accounts for an increase from 29 to 71.5 percent (Tierney, 2012), suggesting that the risk of losing money is more of a motivation than the risk of losing face in front of family and/or friends. From this example it can be concluded that people seem to be willing

to commit to these contracts because they know they cannot commit to their long-term goals on their own. If people were completely rational, they would not want to take on these contracts and risk losing money; they would simply stick to their goals. Unfortunately, people are not rational, and this is why nudges can help them make decisions that are more consistent with their long-term goals.

### 3. Permissibility of nudges in general

This chapter discusses three of the most important moral objections generally raised against nudges and the libertarian paternalistic aspect of this concept.

#### 3.1 Autonomy

An objection often raised against (libertarian) paternalism and the use of nudges is that it impairs autonomy. For this argument, autonomy shall be defined in a broad manner: people's ability to decide for themselves and to act upon these decisions. In using paternalism, the actor – which, in the case at hand, is the government – uses its power (through legislation for example) to protect people from their own imperfect decision-making, thereby impairing their autonomy in order to make them better off as judged by themselves.

Wilson (2011) identifies two main arguments when it comes to paternalism and autonomy. If a certain action was chosen voluntarily, the decision-maker should not be protected from the harm he might do to himself. Apparently, he thought about it and came to a certain conclusion, thereby exercising his autonomy. Next to this, in intervening in the decision-making process, one displays a certain kind of disrespect to the decision-making authority of the person. This is even the case if options are not taken away, as is done in nudging. Feinberg (1986) also states that he considers autonomy to be more important than personal well-being. In fact, he seems to think of it as something of a different level. The well-being is *his* well-being, and therefore he can do with it as he pleases, irrespective of the consequences on well-being, using his autonomy in decisions about it.<sup>2</sup> Moreover, in using nudges, one is taking advantage of the imperfections in human decision-making to alter decisions in a predictable way (Bovens, 2009). This means one does not only disrespect the decision-making authority of the person, but also the person himself. This argument of Bovens can be related to the two levels of nudges: higher and lower order nudges. Lower order nudges aim at intervening at the intuitive level, the level that operates automatically and which the decision-maker has little control over. Using this type of nudges has bigger implications for autonomy than higher-order nudges, which affect the reflective system. Following this line of reasoning, especially taking advantage of imperfections in decision-making seems to involve the lower-order, intuitive, level, making this kind of intrusion more objectionable: the person and his decision-making are not respected.

In line with this argument, Thaler and Sunstein raise the subject of why people cannot just sometimes make the wrong decision, which they call the 'right to be wrong'. This holds in cases in which it truly is a deliberate decision, and not in cases of bad decision-making that result in harmful mortgage choices for example (Thaler & Sunstein, 2008).

---

<sup>2</sup> Feinberg also tries to alter the argument by saying, like Mill does, that autonomous decision making is part of one's well-being (Feinberg, 1986). On the other hand, one can also say that things people rationally prefer may not be preferable from a well-being point of view, but they prefer it nonetheless and want to exercise this preference through autonomous decision-making (Darwall, 2006). The latter argument would make nudges a very strange concept, because they are aimed at increasing well-being by pushing people in a certain direction in decision-making.

This objection is directed at a central part of the concept of nudge, which is influencing people and using their own decision-making imperfections ‘against them’ – in the sense of trying to influence them making use of these imperfections of which they may be unaware – in their own interest. It, therefore, is not an easy one to tackle, because every nudge will try to influence choice, thereby disrespecting the decision-making authority of a person. To what extent people’s ability to decide for themselves – their autonomy – is impaired in the case of health-related nudges, is the topic of the analysis in the next chapter.

### 3.2 Information about preferences

A more practical problem regarding nudges is an informational one. In order to influence decision-making to make people better off *as judged by themselves*, one has to know what people consider better for themselves. As Kant already said:

*I cannot do good to anyone according to my conception of happiness (except to young children and the insane), but only according to that of the one I intend to benefit.* (The metaphysics of morals, 1991, p. 248) (underlining mine)

He displays the concern that one cannot decide for another what is good for him, unless this other person is unable to exercise his or her autonomy in decision-making. Because of this, it is necessary to have information about people’s preferences, so policies can be developed to guide them towards what will make them better off. Now this is where the problem lies – especially at the governmental level: how can one know the preference structure of individuals? According to revealed preferences theory, used by economists, preferences are revealed in the choices people make, in which they show to prefer one option over another (Richter, 1966). However, most of what is observable from outcomes of decision-making is material and measurable, such as income, possessions, or health records – and it is not even sure how the causal connection between choice and outcome is. Immaterial outcomes, like individual levels of happiness, cannot be measured (or not at acceptable cost), and things such as moral principles people have are not easily identifiable in observable outcomes either. This means that preferences cannot be known (at least, to a satisfying degree), and if preferences are not known, changes in choice structures cannot be made in a way that ensures that the final goal is reached: making people better off as judged by themselves.

Moreover, even if preferences *were* revealed, it would not be certain that these preferences are people’s *true* preferences. Due to the possibility of imperfections in decision-making, which means that people are influenced by their choice environment rather than solely using their own preferences as input for decisions, distortions could be caused. This means that there is a possibility of not observing genuine, clean, preferences, and policy-making can therefore not rely on these preferences.

However, even if one *was* to know the exact preference structure of individuals, meaning one would know exactly what to do to make each of them better off, for the concept to be put to practice there should be at least some level of homogeneity in the preference structures of different people. If this is not the case, policy makers cannot adjust the choice structure in a uniform and successful way to guide people to options

that would make them better off. For example, if they want people to be able to keep on living under the same standard after they retire as they do now, then they would very likely nudge people to save accordingly. However, people have different standards of living, so a uniform nudge to saving a certain amount could not be applied. Naturally, some niches based on basic characteristics could exist, meaning that some varieties of choice structures could be set up (different categories of living standards, for example), but extensive fine-tuning would complicate the choice architecture to such a degree that it could bother its users in their decision-making, and would therefore work in the opposite direction.

### 3.3 Truth and transparency

A third objection commonly used against nudges has to do with their form and content. If nudges are applied, we want the information contained in nudges to be true (so, for example, information highlighted – which is the nudge – on food containers given about nutritional values of a product have to be correct), and the presence of nudges to be easily detectible, that is: transparency matters.

As DiSilvestro puts it, transparency is important *“because we do not want to be mere cogs in someone else’s social machine, cut off from the possibility of knowing that we are being nudged and prevented from finding out how and why and by whom we are being nudged”* (2012, p. 14). This, however, does not mean that alarm bells should ring at every instance of a nudge being used, or that it should be stated explicitly on the spot, but people should be able to detect when they are being nudged. Think, for example, of a situation in a supermarket in which the healthier things are more easily reachable, whereas the less healthy products are either on very low or very high shelves. If you wanted to, you could detect this nudge; there is nothing secret about it. Thaler and Sunstein (2008) mention this objection in the context of preventing people with bad intentions from being able to use nudges and to turn them into mere manipulation. Bovens (2009) also stresses that someone paying attention to the scene should be able to see through the intentions of the way choices are structured, so that this person could then decide to intentionally ignore the nudge, or make use of it.

Of course, when nudges are applied by giving information of some kind, this information should be true, so the nudges should represent reality in an accurate way. DiSilvestro claims we do not want to be *‘mere servants of some Platonic noble lie that deceives us for our own benefit’* (2012, p. 14). Even if little lies make us behave in a way that would make us better off, these nudges are still not morally permissible, for the simple reason that they involve incorrect information. If it is the case that this false information, the lie, cannot be detected, manipulation could be possible. Therefore, in order to make decision-making possible, truthful information is required.

The point of this argument is that nudges, contrary to what Thaler and Sunstein proclaim, can be used in a way that is open to malicious intents and manipulation, which would mean lying to people and not respecting them as intelligent human beings. It is therefore that information presented in the choice structures should be true, and nudges towards certain options should be transparent.



An example of this manipulation (albeit a mild form) can be found in a strategy some producers of foodstuffs employ: they label a fruit drink as healthy, because it contains fruits and vegetables, but they do not mention that eating these fruits and vegetables would be even healthier, because a lot of sugar is added to the drink. Other types of drinks that are supposedly good for your bowels, and are therefore claimed to be healthy, may appear to contain other ingredients that turn out to be not so healthy. By labeling their products as healthy, while in fact they are not really, these producers manipulate people in their aim to be healthy. It could lead people to think: I really like this drink (due to the sugar and our taste for it) and it also says it's healthy, so I can take more of it. If they knew sugar was added, however, they would make a different consideration and would be more aware of their intake of the drink. If information stated on the packaging is investigated and turns out to be untrue, these companies can be sued and further measures can be taken. However, not everything can be investigated that easily, so room always exists for these producers to lie in order to promote their product.

## **4. Health-related nudges: what is it that makes health-related nudges a less objectionable type of nudges?**

Given the objections mentioned in the previous chapter, how do these change when we apply nudges to health, being defined as exercise and nutrition? Are there characteristics of health that would make nudges in this context more morally acceptable than in other instances? This will be investigated below, focussing on three main differences that make some of the general objections against nudges less severe.

### **4.1 Relative rather than absolute goal of health-related nudges lessens the informational problem and intrusion**

The first difference focuses on the relative rather than absolute goal of health-related nudges, which reduces operational and conceptual problems with (libertarian) paternalism. This makes the informational and autonomy problem less severe.

The paternalistic aspect in nudge is, according to the initiators of the term, found in its aim of making people better off as judged by themselves (Thaler & Sunstein, 2008, p. 5). This is done by editing the choice architecture in a way that makes it more likely that people choose the option that would make them better off. This paternalistic part of the approach demands an investigation into the preferences of the people being nudged. Therefore, this requires policy makers – *choice architects* in nudge-jargon – to have information about these preferences, and to furthermore take them as a starting point for developing the choice structure in a way that makes it more likely that these preferences are satisfied compared to the status quo. In the case of the government applying nudges, policy-making and implementation would come down to the government adapting the choice architecture and using this to nudge every inhabitant faced with that choice.

These requirements would pose a real problem in contexts in which nudges are usually put to practice, because preferences are not known and/or not sufficiently homogeneous. A good example of this is saving behaviour. Many people structurally save too little to be able to maintain their standard of living after their retirement and are therefore nudged in different ways to save more for that time in their lives (Thaler & Sunstein, 2008). However, different people have different preferences about how to intertemporally divide their money – between ‘now’ and ‘later’ – depending on character traits, possibilities and financial endowments, making it difficult to implement a general choice structure in which nudges are incorporated.

In the context of health-related nudges, however, this problem is far less severe due to the relative rather than absolute goal of policy makers. This is the case because, to a certain extent, relative preferences concerning health can be assumed to be homogeneous: every person wants to be healthy and would rather have *more* of health than less, other things being equal. If the preferences about health were considered in their absolute form, then policy makers would have to know exactly how much health and what kinds of it every person

would want to have. However, since relative preferences are the only requirement for applying health nudges, solely preferences about wanting to have more or less health have to be known. And, as said before, these preferences can be assumed to be homogeneous, because roughly everyone would rather be more healthy than less. This claim is also supported by various theories, based on consensus, about minimum standards of living all human beings should have. For example, Martha Nussbaum mentions bodily health as one of the central human functional capabilities (Nussbaum, 1999). These capabilities are prerequisites for flourishing as a human being, and should therefore be guaranteed by every government.<sup>3</sup> This line of reasoning shows that the informational problem is not significant in the case of health nudges. Because of the relative homogeneity, policy makers do not have to ‘guess’ people’s preferences concerning health, nor do they have to make too strong assumptions about them. In other instances of nudge, for example in the case of energy consumption and saving, it could also be that people would generally rather save more than less or consume less energy than more, but this preference is not as basic as the desire to be healthy, since that seems to be prior to wanting anything else. This can also be brought back to the central human functional capabilities, in which health plays a big role, but material aspects hardly do (Nussbaum, 1999). Additionally, since the aim of policy makers seems to be more relative in the case of health than in other cases – they just want to make people act more healthily, which is relative rather than absolute –, the exact preference ordering that takes into account all different alternatives, as well as how health preferences relate to other preferences within each individual, do not have to be known in order for them to be able to nudge people, which lessens the problem of implementation.

This homogeneity in preferences about relative health can also be seen as something that would make nudges in health more morally permissible than in the case of organ donation, for example, since it is thought few people would *object* to being led to healthier rather than less healthy behaviour, thereby making the nudge less intrusive. However, since organ donation is a subject many people consider *queasy*, as mentioned in section 3.1, some people do not want to be confronted with it unless they decide so themselves, making this application less morally permissible than health nudges. The relative absence of objections to nudges also applies to other applications than health though. An example of this is the case of pension savings: every person would rather have more money to spend now than later, given time preferences, but they do know that they should save for later. Therefore, any help they get is welcome in making better the intertemporal money decisions, and since the nudge is avoidable, this will evoke little objection.

The libertarian component of nudges furthermore implies that choice options are not taken away and that the ‘good’ option can still be avoided at low cost, implying that the original choice set is preserved. However, due to the application of nudge, the choice set is certainly altered in a way that puts one of the options forward as the one preferred by the policy makers. These choice architects will try to influence decision-making so that eventually this option is chosen. This can be visualised and explained in the following way. Suppose the choice at hand has five options to choose from, of which one option is the one choice architects

---

<sup>3</sup> Whereas Nussbaum considers this list to be objective in her early work, in her later work she argues that the “capabilities can be the object of an *overlapping consensus* among people who otherwise have very different comprehensive conceptions of the good” (Nussbaum, 2000, p. 5), which serves as a justification for the items on the list.

want people to choose. The dots below represent the five options, and the bigger the dot, the bigger the nudge towards it.



Looking at this choice set in which people are nudged, the choice set is preserved, but one of the options is highlighted. Presented with the choice, people will be more likely to choose the well-presented option. Therefore, even though all options are still available, it is still thought that this specific option is pushed forward, thereby influencing the choice and impairing people's autonomy. An example of this situation is when *default options* or *subconscious priming* are used, thereby leading people to this option and requiring extra effort to choose a different one.

In the context of health-related nudges, however, due to the relative rather than absolute goal of the nudge, the choice set is not manipulated in the way described above. This is the case because the aim here is to make people live more healthily, not to reach a certain level of health – if that is possible to identify at all. All we need for this is to nudge them to make better decisions so they will exercise more and eat more healthily, but there is no specific quantity attached to this. Even though there is an ample array of health guidelines, this is not ultimately what we strive for using these nudges. This quantitative approach is present, however in other cases of nudge, such as the savings example, and the absolute approach is the case in organ donation: the goal is to get people to choose 'yes', so the absolute goal is the starting point. In the health case, however, the goal is comparative: it aims at making people live more healthily, not to have them walk for half an hour and eat an apple, kiwi, banana and five ounces of vegetables a day. These results would be nice if they did not do so before, but not what is strictly intended. This means that, since more than one option in the choice set is preferred to the status quo, rather than having just one option that is advocated by policy makers, there is less interference in individual choices. Visualised:



This is a situation in which there is a least-preferred option, which may very well be the status quo, and the options to the left of it are preferred. For example, if the choice is about what to have for lunch in a highway restaurant, the smallest dot could represent a burger menu, every dot to the left is a healthier option, and the biggest could be a well-balanced meal with a big portion of vegetables. Using nudges this way, by increasingly highlighting better options, or highlighting all options except the frequently chosen worst option, means that not just one option is preferable, but all options that are better than the one people normally choose. If people are presented with a choice structure in which one of the choices is put forward as the ideal option, autonomy would be impaired more than in a case in which all options that are better than the status quo are highlighted. Therefore, given that the health-related nudge at hand is aimed at making people live *more* healthily, the distinction between absolute and relative goals makes the conceptual problem of paternalism of the impairment of autonomy less severe in this health case relative to other applications of nudge, such as organ donation.

There are, however, also other applications of nudges in which this relative choice could be applied, meaning that health nudges are not unique in this. For example, in energy consumption the aim is to get people to consume less energy, so the relative component is present. Actually, this is already visible in the case of buying a washing machine: the machines have labels, and the more environmentally friendly the machine is, the greener and bigger the logo, making these machines more likely to be bought by customers. Therefore, any area of applying the concept of nudge that makes it possible to add this relativity in the choice framework, as is the case with health nudges, would be more morally permissible than making one of the choice options the only highlighted option.

This relative approach in which people are pushed towards any alternative that is better than their current option, could, however, have negative side-effects: nudges could lose their effectiveness. Being presented with many options that are supposedly preferable to their status quo – although in different degrees – could cause people to be unable to see the wood for the trees, and just, perhaps even randomly, picking an option.

Although it would already be an improvement if people chose *a* better option, it would be best if they chose *the best* option. From an effectiveness point of view, it could be argued that promoting one option as the best would be more desirable than promoting several options, in different strength, as better options. As discussed, though, from a moral point of view this would not be desirable, since this would make the nudge more intrusive.

## **4.2 Habits are more easily created in the health context, leading to more intrinsic motivation to stick to long-term goals**

The second difference that makes health-related nudges less objectionable is that they are more likely to lead to real, long-term changes, because the nature of these nudges makes it more likely that people form a habit.

In his work on the ethics of nudge, Luc Bovens (2009) states that nudges, in their aim at guiding people towards options that make them better off, are not likely to build moral character. He explains that our inability to act in accordance with our overall preference structure is caused by a lack of self-control, which will not be solved if we are nudged towards a better alternative. There would be short-term effects, but once people get used to these nudges, more severe ones will be necessary to have the same effect. He also expresses his concerns for a slippery slope causing people give away responsibility over their own well-being to an external authority, which he calls *infantilisation*. This can be interpreted as an impairment of autonomy.

In the case of health-related nudges, there are two arguments against the problems addressed by Bovens. The first is directed at the lack of self-control, and the second is about habit formation, both specifically in the light of health-related nudging. The former explains that the lack of self-control is caused by visceral influences, which play a large role in health behaviour, thereby making it less likely that moral character is built in health-related decision-making than in other choice contexts in which nudges are applied. The latter shows that the nature of health makes it more likely that the behaviour resulting from a nudge is turned into a habit, internalising a certain preference for the better option.

Moral character can be described as personal convictions about right and wrong, of which you have a disposition to apply in the same way in different situations. When it comes to health nudging, this would be translated into what is good for your health and what is not, or at least in a relative sense: which option would be better and which would be worse. When Bovens talks about a lack of self-control and takes nudges as an instrument for trying to correct for it, building of moral character can be undermined if people are less inclined to display their convictions in their actions if the nudges would suddenly be absent; he therefore argues that nudges do not build moral character.

The lack of self-control, or rather the weakening of self-control, is a phenomenon thoroughly investigated by behavioural economist George Loewenstein. In an article written in 1996, he points out an anomaly in decision theory: visceral factors are not taken into account well enough to be able to explain certain types of behaviour. According to Loewenstein, visceral influences “include drive states, such as hunger, thirst and sexual desire, moods and emotions, physical pain, and craving for a drug one is addicted to” (Loewenstein, 1996, p. 272). Due to the bodily nature of health, these influences are likely to play a larger role in the case of health-related goals than in other goals towards which people could be nudged, such as saving money or energy, which have mainly monetary and social consequences. It is in the face of these visceral influences that this so-called lack of self-control arises, while the conviction about the right thing to do to establish long-term goals and the disposition to behave accordingly are still present. The difference

with other applications of nudges is that in these other types, visceral factors only play an indirect role, making them less important to self-control. In the case of saving behaviour, this indirect role means that saving behaviour does not immediately entail physical sacrifices or sensations, not until people for example need to replace their cars with bicycles or lower their food intake in order to be able to save money, whereas health behaviour has immediate physical effects. Therefore, the lack of self-control in these other contexts is more likely to be caused by absence of the disposition to behave according to a certain conviction. So in health, the disposition can be present, but visceral factors may directly prevent people from acting on their convictions. An example of this is the case in which someone wants to eat less chocolate, has the disposition to behave accordingly, but when presented with chocolate at a birthday party, visceral factors of appetite and smell can come into play, causing this person to eat the chocolate anyway.

Later on in his work, Loewenstein specifies the underlying concept of self-control as willpower (Loewenstein, 2000). The idea is that willpower can be compared to a muscle: you can put it to use, but only to a certain extent. Once this limit has been reached, it is extremely hard to stick to your originally well-intended plan of approach, and visceral influences take charge of your decision-making, at least to a certain extent. The problem is that individuals systematically overestimate the strength of their willpower, for example to make the healthy (i.e. the right) decision. This is caused both by the fact that they underestimate the influence of desires and visceral factors on future behaviour and by their unawareness of the decline in willpower as it has been put to use for a longer period of time. Since willpower is often put to use in working towards a long-term goal, it can be very challenging to keep it up, especially when you are exposed to situations in which you initially thought you would be ‘strong’. If you eventually give in to your desires as your willpower drops below a certain level, this does not mean, however, that you don’t have that long term goal in mind along the way, you may just not always be able to act accordingly.

Since the health-related decisions considered are very prone to visceral influences, it is very unlikely that people will build more character – in the sense of having the disposition to apply your convictions, so actually sticking to your goals – without nudging them towards healthier options. This can be seen in the fact that people who want to stick to their diet are very prone to moments of weakness when presented with their ‘guilty pleasures’. If they love ice cream for example and all of their friends are having some. This creates the appetite – the visceral influence – and makes it more difficult to stick to long term goals. Another type of example is that people who get home from work feel tired – a visceral factor – and do not want to go to the gym, even if they know they will feel much better afterwards. Nudges can help in these circumstances. Because of the previous explanation and examples, the argument that nudges do not build moral character is less of an argument against using nudges in the health-related context, because people are very likely to build moral character in this setting anyway. Very direct visceral influences that are present in the health context lead to less self-control, as opposed to other areas of application where visceral influences play a more indirect role.

Bovens also warns for *infantilisation* of the people being nudged, meaning a decreased responsibility for one's own welfare. This decreased responsibility is caused by the fact that nudges form external cues to making a decision in a certain way. Bovens' fear is that people will no longer think about the choice, and just choose the option towards which they are nudged, meaning there cannot be a long-term effect if the nudges are stopped.

This argument, however, can be countered by the process of habit formation, which especially applies to the case of health-related nudges due to two factors. Firstly, as put forward on the previous page, the nature of health makes visceral factors especially influential in decision-making. Since the nudges are meant to make people decide more healthily, their health will improve, cause them to feel better, which is a physical, visceral effect. Secondly, the health decisions at hand – namely the ones that have to do with nutrition and exercise behaviour – are more likely to have short-term effects, thereby providing short-term feedback to the decision makers. For example, if people start exercising more regularly, they will start to feel better in a matter of days or weeks. These two factors combined mean that the decision-making that is influenced by nudges will have short-term visceral effects: the rewards are visceral and placed earlier in time, making it easier to create a habit than in other types of nudges, such as pension savings, donor registrations, or energy consumption. In pension savings and donor registration, the effects will be long term; in fact, in donor registration you will not even experience the actual effect, but that is not the point. With the savings, there is also no direct visceral reward present. As for the energy consumption nudge, the feedback will be short-term: people can see how much energy they have used and will see this on their bills. However, no direct visceral factors are involved. Therefore, in nudge cases that do not have short-term and/or visceral effects, habits are less likely to be created compared to nudges in health decisions.

Once the healthier habit has been formed, people will be more likely to sustain this healthy behaviour, even if the nudges are stopped. This is the case because they incorporate it into their daily lives, thereby acting out their original (long-term) preferences to live more healthily, or even changing them towards more healthy preferences.<sup>4</sup> It has also been shown that habits (which start out with reasoned action, but eventually involve a non-deliberative process) form a good predictor of actual exercise behaviour, whereas merely having the intention to exercise (which is a deliberate process and has more to do with long-term goals) is not informative for predicting this behaviour at all (Aarts, Paulussen, & Schaalma, 1997).

Therefore, the objection raised against nudges that it might lead to infantilisation is less strong in the case of health-related nudges, thereby making the autonomy problem less severe. Because by being nudged to live more healthily, this way of living is more likely to be internalised through the process of habit formation, and therefore people will eventually make decisions based on their resulting changed or reinforced preferences, rather than merely being pushed in a certain direction by external authorities. This solves the autonomy problem, at least partly, since the influencing only takes place at the beginning of the nudging process.

---

<sup>4</sup> For an explanation of habit formation models in the context of intertemporal choice, see Wilkinson, N. (2008). *An introduction to behavioral economics*. Basingstoke: Palgrave Macmillan, p. 237.



A possible objection Bovens could make to the latter argument is that habits are not genuine deliberative processes and therefore creating habits does not involve more autonomy than infantilisation would. In psychology, habits are seen as relatively simple acts that are frequently or customarily performed without much deliberation (Mixon, 1980; Verplanken & Aarts, 1999). Therefore, habits are actions that you perform just because you are used to doing them in a certain way, and could therefore be seen as thoughtless repetitions.

However, *adopting* the habit still is something the person him- or herself has control over. It appears that habits are formed because of a cognitive orientation towards a certain goal, called goal-directed automatic behaviour, and this orientation does persist in later repetitions, even though the deliberation is not a part of the habitual acts (Verplanken & Aarts, 1999; Aarts & Dijksterhuis, 2000). If the nudge disappears, it could be that people who only bought healthy food because of the relative cheapness of it, stop buying it, whereas people who have come to like it and created a habit, will continue to buy the food even when the nudge has been stopped, because the goal orientation remains: people aim at living healthier lives. Therefore, acting out the habit, like all other routine actions, does not immediately involve autonomy in the form of deliberation, but choosing to adopt the action and create the habit does.

### **4.3 Rather than impairing it, health nudges may even save autonomy**

The general autonomy problem still holds partly in the case of health-related nudges, since people are being influenced in their decision-making, which means that their decision-making authority is not fully respected and their autonomy is impaired. This influence takes place through an exploitation of the imperfections in their decision-making and manifests itself in guiding people in a direction – the healthier one – rather than deleting options from their choice list; people can still choose the unhealthy options. However, there are some aspects of health nudges related to autonomy that can be used to argue in favour of health nudges as opposed to other instances of nudge.

First of all, in the case of health, it is questionable whether people should be given full autonomy and not be influenced by external authorities in their decision-making about health behaviour at all. Gaylin (1996) raises the issue of ‘the perversion of autonomy’ and the threat that lies in worshipping it. Promoting autonomy, Gaylin says, is based on the idea that people are rational and are able to decide for themselves what is best for them. Drawing on the subject of the lack of self-control discussed above however, people are not fully rational and emotional and physical circumstances have a large influence on decision-making, especially in the case of health. For example, people are more likely to eat comfort food when they are feeling down, but it can also be the case that good looking people who are eating unhealthy food makes it more likely that they themselves eat more unhealthy food. There is ample evidence of people actually harming themselves in this ‘free decision-making’ setting, ending up with severe and even deadly diseases. This consequence could lead to the conclusion that health is more important than autonomy about health. Health could actually be considered a prerequisite for autonomy, since autonomy implies being *able* – mentally and physically

capable – to decide for yourself and act upon your decisions. Some guidance in the form of adapting the choice architecture may therefore be more permissible in the case of health than in other applications of nudge, in which people show less self-destructive behaviour when left completely free to choose. For example, consuming too much energy will not have a direct effect on you as a person – these effects are postponed. The case of pension savings is more direct than this, but also only causes indirect harm by for example not having enough money for food and health care in old age and having to rely on others or adjust behaviour drastically. In this case, wrong decision-making, if left alone, can also have severe consequences for a person’s autonomy. However, the effect on autonomy seems to be less direct and/or severe in a physical sense than in the case of bad decision-making in health.

If one, however, does consider autonomy to be more important than health, another consideration pops up. A subject to be raised in the context of health-related nudges on a public scale is whether every aspect of nudges in public policy should be seen as objectionable because of the paternalistic and autonomy impairing aspects to it. Wilson (2011) claims that public policy is bound to be paternalistic because interfering in autonomy and lack of individual consent are endemic to that context, but that does not mean it is not justifiable, even though it is paternalistic. Therefore, about each policy it should be asked whether the infringements of autonomy are justifiable. Given that health is directly important for a population and a country to fare well, policies aimed at enhancing health would be more justifiable, according to this reasoning, than other cases of nudge, such as pension savings, which is more indirectly relevant to the welfare of a country. This also has to do with the idea of the government having to take up on its responsibility for good governance, which does not automatically make every policy design merely a paternalistic manipulation of behaviour (Carter & Hall, 2012).

Approaching the autonomy problem from a different point of view, namely by looking at the actual consequences rather than the concept, using nudges could actually lead to an increase in people’s autonomy. Daniels (2008) stresses in a book called *‘Just health’* that social circumstances that prevent people from being able to make decisions or to act out these choices should be taken care of and solved. In this context he talks about people’s opportunity range, which is understood to be the “array of life plans reasonable persons are likely to develop for themselves” (p. 44). According to Daniels, people should be able to act out these life plans without any impairment by health.

When the use and the non-use of health-related nudges are contrasted, not using them makes diseases like obesity and others, which can be prevented by the right nutrition and enough exercise, more likely to occur than when nudges are used. As a consequence, if people keep on living the way they do now, statistics like those in the United States showing that roughly thirty percent of children between 9 and 12 and thirty-five percent of adults are overweight or obese will not be out of the ordinary (American Heart Association, 2013). Due to health-behaviour related diseases, these people become immobile or may even be affected mentally. Immobility implies not being able to do however you please, and can also mean that other people may have to start making decisions for you. Mental affection due to such diseases can take on many forms.

Obesity has been shown to be associated with significant increases in lifetime diagnosis of major depression, bipolar disorder, and panic disorder or agoraphobia (Simon et al., 2006), meaning that those affected cannot function normally anymore. Consequently, individuals choosing unhealthy alternatives and eventually developing these health-behaviour-related diseases, become impaired to make decisions for themselves.

This impairment can also take place in other examples of nudges, generally in all applications in which people's own decision-making would put them in a position in which they are very limited in their freedom to decide. The savings example could, once again, be an interesting one to clarify this claim. If people save too little, they could put themselves in a position in which they have very little room to decide, given their financial circumstances. It could, for example, be the case that they become financially dependent on other people, limiting their options regarding spending, or any other kind of limitation on their decision-making. However, the most important thing to stress here is the impairment in this case is not primarily bodily – physical or mental – but external, as is the case with health nudges; in the savings example the impairment is mainly externally caused by limitations. Considering nudges regarding energy consumption, this impairment of autonomy would only be as severe as in the health case if the environment were so extremely polluted that toxic gasses for example make it impossible for people to think clearly or become physically disabled. However, people's individual behaviour over their lifetimes is not likely to cause these kinds of terrible effects, and can therefore not be prevented by nudging these individuals towards consuming less energy. Considering both of these examples, even though nudges in decision-making about savings could lessen the impairment of autonomy in an indirect, external sense, health-related nudges can be taken to contribute more to autonomy, because they enhance people's physical and mental ability to make decisions.

Furthermore, it can be argued that this impairment of autonomy, caused by unhealthier decision-making in the absence of health nudges, is even more severe than in the case of nudging people towards healthier behaviour. If nudges are used, people who currently have this problem can be helped and become more capable of making decisions. So, especially in the case of health, something you should have to function normally – or rather: the lack of which causes an impairment of autonomy – nudging can be said to be more beneficial to autonomy than not nudging.

## 5. Conclusion

The objective of this thesis was to show how the objections usually raised against nudges change when they are applied in a health context. For this purpose, firstly, the concept of nudge was explained in the context of libertarian paternalism. Nudges are aimed at making people better off as judged by themselves, and can be defined as “ways of influencing choice without limiting the choice set or making alternatives appreciably more costly in terms of time, trouble, social sanctions, and so forth. They are called for because of flaws in individual decision-making, and they work by making use of those flaws” (Hausman & Welch, 2010, p. 126), to which should be added that manipulation is not considered a nudge, and the mechanism should have unconscious influence. Next, various internal agency problems which nudges try to correct were discussed, as well as the forms the nudge can take on and various areas of application. After having defined health for the purpose at hand as the amount of exercise and healthy nutrition, possible applications of nudge on these determinants were put forward.

In the first part of the main body of the thesis, three of the most important objections against nudges in general were mentioned. A first objection lies in that nudges impair the autonomy of the nudgees, because they are not left free in their decision-making and their decision-making authority is not respected. The second argument regards the lack of information about preferences and the heterogeneity of preferences: people’s preference structures are usually not known and even if they were, they would differ so much across the population that it would be difficult for policy makers to design general and successful policy based on them. The third objection raised against nudges is more of a requirement that nudges have to meet in order to be permissible, which is that they should be transparent and should convey truthful messages: people cannot be lied to just to get them to alter their decision-making, and someone who would want to, should be able to detect if and how he/she is being nudged in a certain situation.

Turning to health-related nudges, the objections raised against nudges in general change due to the nature of health. The first important aspect of change is that health nudges have a relative rather than an absolute goal: policy makers try to make people behave *more* healthily (exercise more often and/or eat more healthily) rather than setting a specific goal they must reach. This means not just one choice alternative is preferred by policy makers, but all the options that are better than the worst one or the status quo are considered improvements and therefore favourable. Another implication is that there is less infringement of individual autonomous decision-making, since people can choose many alternatives and still behave in a way that is better for them, rather than just one option being the better option. Furthermore, because health is something every person is assumed to want more of rather than less, the lack of information about true preferences is less severe, and preferences are taken to be more homogeneous than in other applications of nudge, such as organ donation: everyone would rather be healthier than less healthy, but not everyone wants to be a donor.

The second change in the objection regards moral character and long-term effects. Due to the physical nature of health and the consequences of behaviour being more visible in the short term, habits are more easily created and sustained, because of an intrinsic motivation in the form of preferences to maintain

the healthy behaviour. Nudges will therefore not lead to infantilisation, a concern Luc Bovens raises, meaning that people are no longer capable of autonomously making the right decision. Nudges will first be influenced solely by choice structure, but once the habit is created, they will develop or reinforce the preference for the healthy option over the less healthy option, and then make better decisions autonomously based on their preferences rather than on the nudge. Also, the determinants of health at hand – exercise and nutrition – are influenced directly by visceral factors such as hunger, and therefore require a lot of willpower. Since willpower is thought to be like a muscle and cannot be exercised easily over a longer period of time or in times of temptation, it can be argued that nudges are especially valuable in the case of health. This is the case because the lack of self-control is more imminent than in other cases of nudge, such as savings, where visceral factors only play an indirect role in behaviour.

Lastly, the autonomy objection also changes at least partly in the case of health-related nudges. As is mentioned in the previous paragraph, nudges impair people's autonomy because they are influenced through making use of imperfections in their decision-making. However, in the case of health, it is questionable whether people should really have full autonomy in decisions about it: when it comes to health, people actually display self-harming behaviour in case of free decision-making. Health may be more important than autonomy. On the other hand, if autonomy is considered more important than health, the fact that we are looking at the governmental level entails public policy, which is bound to be at least partially paternalistic. Therefore, on a case-by-case basis, policies should be investigated on their infringements of autonomy, so see whether they are justifiable. Another approach to autonomy in relation to health is looking at the practical consequences, rather than the theoretical concept. Since people who are not nudged may be more likely to display unhealthy behaviour, the chances of developing behaviourally caused diseases, such as obesity, are higher. The consequences of these diseases may be that they are no longer capable of functioning on their own and require help. In that case, they can no longer fully decide for themselves and therefore lose some of their autonomy, by actually losing certain options that were previously available or no longer being physically or mentally able to make decisions. It can be argued that this impairment on autonomy is even more severe than in the case of nudging people towards healthier behaviour. So, especially in the case of health, something you should have to function normally – or rather: the lack of which causes an impairment on your autonomy –, nudging can be said to be more beneficial to autonomy than no nudging.

Of the objections raised against nudges, the one about truth and transparency still holds. However, the concern about the misuse of nudges in case of health and in hands of the government can be considered less imminent, especially if nudges are transparent.

The overall argument of this thesis can be summarised as follows: a number of particular characteristics of health nudges, such as their relative rather than absolute goals, their role in habit formation, the influence of visceral factors on decision-making in health and the direct impact health improvement has on a person's autonomy, make them less vulnerable to the most prominent objections raised against nudge. Even though other non-health nudges share some of these same characteristics, health nudges are unique in their combination of them and can thus be considered more morally permissible.

Though the most important and commonly raised concerns about the use of nudges were discussed, not all possible objections were explored in this thesis. Mainly the problem of implementation plays a big role (how nudges should be shaped and used to be effective), but also concerns about the concept of nudge itself, such as the question how the point of autonomy and self-harm should be balanced out.

The goal of the thesis has been reached, namely showing that the objections do change when we turn to health-related nudges, and that they actually become less severe. However, since the list of objections put forward was not comprehensive, it could be the case – although it seems very unlikely – that there are objections that actually get more serious when we turn to health-related application of nudges. Further research could investigate whether this is the case, and maybe even show new objections that could arise when turning to a health setting.

## 6. Literature

### Articles

- Aarts, H., Paulussen, T., & Schaalma, H. (1997). Physical exercise habit: on the conceptualization and formation of habitual health behaviours. *Health Education Research*, 12(3), pp. 363-374.
- Aarts, H., & Dijksterhuis, A. (2000). Habits as knowledge structures: automaticity in goal-directed behavior. *Journal of Personality and Social Psychology*, 78(1), pp. 53-63.
- Abouna, G.M. (2008). Organ shortage crisis: problems and possible solutions. *Transplant Proceedings*, 40(1), pp. 34-38.
- Blumenthal-Barby, J. S., & Burroughs, H. (2012). Seeking better health care outcomes; The ethics of using the nudge. *The American Journal of Bioethics, AJOB*, 12(2), 1-10.
- Carter, A., & Hall, W. (2012). Avoiding Selective Ethical Objections to Nudges. *The American Journal of Bioethics*, 12(2), pp. 12-14. doi: 10.1080/15265161.2011.634488
- Castelo, N., Reiner, P.B., & Felsen, G. (2012). Balancing Autonomy and Decisional Enhancement: An Evidence-Based Approach. *The American Journal of Bioethics*, 12(2), pp. 30-31. doi: 10.1080/15265161.2011.634951
- Darwall, S. (2006). The Value of Autonomy and Autonomy of the Will. *Ethics*, 116, pp. 263–284.
- DiSilvestro, R. (2012). What Does Not Budge for Any Nudge? *The American Journal of Bioethics*, 12(2), pp. 14-15. doi: 10.1080/15265161.2011.634956
- Gaylin, W. (1996). Worshiping autonomy. *The Hastings Center Report*, 26(6), 43-45.
- Gilligham, K., & K. L. Palmer (2013). Bridging the Energy Efficiency Gap: Insights for Policy from Economic Theory and Empirical Analysis (January 25, 2013). Resources for the Future Discussion Paper No. 13-02. doi: 10.2139/ssrn.2206995
- Hausman, D. M., & B. Welch (2010). Debate: To Nudge or Not to Nudge. *Journal of Political Philosophy*, 18(1), pp 123-136.
- Loewenstein, G. (1996). Out of Control: Visceral Influences on Behavior. *Organizational Behavior and Human Decision Processes*, 65(3), pp. 272-292.
- Loewenstein, G. (2000). Willpower: A decision-theorist's perspective. *Law and Philosophy*, 19, 51-76.
- Mitchell, J. B., & A. J. Moore (2011). Lifetime individual retirement arrangements: An application of

- Thaler and Sunstein's nudge. *Journal of Accounting and Finance*, 11(2), pp. 16-25.
- Mixon, D. (1980). The place of habit in the control of action. *Journal for the Theory of Social Behaviour*, 10(3), 169-186.
- Plante, T. G., & J. Rodin (1990). Physical Fitness and Enhanced Psychological Health. *Current Psychology: Research & Reviews*, 9(1), pp. 3-24.
- Richter, M. K. (1966). Revealed Preference Theory. *Econometrica*, 34(3), pp. 635-645.
- Shafer-Landau, R. (2005). Liberalism and paternalism. *Legal Theory*, 11(3), 169-191. doi:10.1017/S1352325205050081
- Simon, G.E., Von Korff, M., Saunders, K., Miglioretti, D.L., Crane, P.K., van Belle, G., & Kessler, R.C. (2006). Association between obesity and psychiatric disorders in the US adult population. *Archives of General Psychiatry*, 63(7), pp. 824-830.
- Verplanken, B., & Aarts, H. (1999). Habit, Attitude, and Planned Behaviour: Is Habit an Empty Construct or an Interesting Case of Goal-directed Automaticity? *European Review of Social Psychology*, 10(1), pp. 101-134.
- Waller, K. V., & R. C. Bates (1992). Health Locus of Control and Self-Efficacy Beliefs in a Healthy Elderly Sample. *American Journal of Health Promotion*, 6(4), pp. 302-309.
- Wilson, J. (2011). Why it's time to stop worrying about paternalism in public health policy. *Public Health Ethics*, 4(3), pp. 269–279. doi: 10.1093/phe/phr028
- Yun, D., & Silk, K. J. (2011). Social norms, self-identity, and attention to social comparison information in the context of exercise and healthy diet behavior. *Health Communication*, 26(3), 275-285. doi: 10.1080/10410236.2010.549814

## **Books**

- Daniels, N. (2008). *Just Health: Meeting Health Needs Fairly*. New York: Cambridge University Press.
- Feinberg, J. (1986). *The Moral Limits of the Criminal Law: Volume 2. Harm to Self*. Oxford: Oxford University Press.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. New York: Farrar, Strauss, Giroux.
- Kant, I. (1991). *The Metaphysics of Morals*. Cambridge University Press.



Nussbaum, M. C. (1999). *Sex and Social Justice*. Oxford University Press.

Nussbaum, M. C. (2000). *Women and Human Development: The Capabilities Approach*. Cambridge University Press.

Thaler, R. H., & C. R. Sunstein (2008). *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Yale University Press, 2008.

Wilkinson, N. (2008). *An introduction to behavioral economics*. Basingstoke: Palgrave Macmillan.

## **Editorials**

Bovens, L. (2009). The Ethics of Nudge. In T. Grüne-Yanoff, & S.O. Hansson, *Preference Change: approaches from philosophy, economics and psychology*. Springer, pp. 207-220.

## **Government publications**

Behavioural Insights Team (2010). *Who we are*. Retrieved on May 5<sup>th</sup> 2013 from <https://www.gov.uk/government/organisations/behavioural-insights-team>.

Hoeymans, N., Melse, J.M., & Schoemaker, C.G. (2010). *Gezondheid en determinanten*. (Deelrapport van de Volksgezondheid Toekomst Verkenning 2010: Van gezond naar beter.) RIVM-rapport nr. 270061006. Bilthoven: RIVM, 2010. Retrieved on June 12<sup>th</sup> 2012 from <http://www.rivm.nl/bibliotheek/rapporten/270061006.pdf>.

## **Theses**

Rentier, C. R. (2012). *Op weg naar consistentere en betere individuele gezondheidskeuzes met behulp van gedragseconomische inzichten*. (Unpublished thesis). Erasmus University Rotterdam, Erasmus School of Economics. Retrieved from <http://hdl.handle.net/2105/12487>.

## **Internet resources**

American Heart Association (2013). Retrieved on April 22<sup>nd</sup> 2013 from [http://www.heart.org/idc/groups/heart-public/@wcm/@sop/@smd/documents/downloadable/ucm\\_449855.pdf](http://www.heart.org/idc/groups/heart-public/@wcm/@sop/@smd/documents/downloadable/ucm_449855.pdf)

Donate Life New York (2013). *All About Donation: Organ Donation Statistics*. Retrieved on April 28<sup>th</sup> 2013 from <http://www.donatelifeny.org/about-donation/data/#Data US1>.

StickK.com (2013). *About stickK.com*. Retrieved on March 27<sup>th</sup> 2013 from [www.stickk.com/about.php](http://www.stickk.com/about.php).

## **Newspaper articles**

Bates, D. (2012, August 4). Obama adviser and 'Nudge' author Cass Sunstein to leave White House. *The Telegraph*. Retrieved on May 5<sup>th</sup> 2013 from <http://www.telegraph.co.uk/news/worldnews/barackobama/9451997/Obama-adviser-and-Nudge-author-Cass-Sunstein-to-leave-White-House.html>.

Tierney, J. (2012, January 8). Be It Resolved. *The New York Times*, p. SR1.