Under What Conditions, If at All, Can (Psychological) Strategic Behavioural Influences Be Justifiably Used to Shape People’s Choices?

Eleanor Whitehead

Supervisor: Elin Palm
Examiner: Lars Lindblom
ABSTRACT

The publication and mass appeal of Richard Thaler and Cass Sunstein’s book *Nudge: Improving Decisions about Health, Wealth and Happiness (Nudge)*, in 2008, illuminated behavioural economics, in the public and political domain. Nudging, a technique derived from behavioural economics, offers a fresh element to the long-time debate between paternalism and freedom, since proponents believe it can simultaneously preserve freedom of choice and serve as a means to influence behaviour. Unsurprisingly, in the decade or so since *Nudge*, private corporations and governments alike have shown great interest in the behavioural steering techniques derived from behavioural economics. This thesis explores the ethical implications and the various means by which governments and the private sector influence behaviour, specifically individual decision making. Since many of the methods overlap in purpose and practice, I make distinct three techniques: nudging, boosting and market advertising. These steering techniques range from transparent and educative to sub-conscious and manipulative methods; as such ethical justification for their employment varies. This thesis concludes by stating transparency as a condition for ethical behavioural influencing since non-transparent or covert methods do not uphold true freedom of choice. Furthermore, the implementation of non-transparent influences carries the potential for further violations of individual autonomy.

“Because to take away a man’s choice, even his freedom to make the wrong choice, is to manipulate him as though he were a puppet and not a person.” – Madeline L’Engle
Contents

1. Introduction .................................................................................................................................. 4
2. A brief background in behavioural insights and choice steering ................................................. 5
   2.1 Two systems............................................................................................................................... 6
   2.2 Humans and Homo economicus ............................................................................................. 7
   2.3 Attitude and decision making .............................................................................................. 9
   2.4 Heuristics or Rules of thumb .............................................................................................. 10
3. What differences in ethical challenges do nudges, boosts and market advertising face? .......... 10
   3.1 What is a nudge? .................................................................................................................... 11
       3.1.1 Libertarian Paternalism ................................................................................................. 11
       3.1.2 ‘Fuzzy’ or ‘mistaken’ nudges ....................................................................................... 12
       3.1.3 A technical definition ................................................................................................. 13
       3.1.4 A loose definition of nudging ..................................................................................... 13
   3.2 How do nudges differ from boosts? ....................................................................................... 14
       3.2.1 Intervention target ......................................................................................................... 14
       3.2.2 Roots in research programs and evidence .................................................................. 15
       3.2.3 Causal pathways ......................................................................................................... 16
       3.2.4 Normative implications .............................................................................................. 17
       3.2.5 Assumptions about cognitive architecture ................................................................ 18
       3.2.6 Empirical distinction criterion (reversibility) .............................................................. 18
       3.2.7 Programmatic ambitions ............................................................................................ 19
   3.3 Market advertising and nudge .............................................................................................. 19
       3.3.1 Purpose .......................................................................................................................... 20
       3.3.2 Accountability ............................................................................................................... 21
       3.3.3 Roots in research and methods of implementation ....................................................... 22
4. Justifying or rejecting (psychological) strategic behavioural influences ................................... 23
   4.1 Sustainability .......................................................................................................................... 24
   4.2 Choice Architects .................................................................................................................. 25
       4.2.1 Manipulation .................................................................................................................. 25
       4.2.2 Corruption .................................................................................................................... 26
       4.2.3 Control .......................................................................................................................... 26
   4.3 Where’s the harm? .................................................................................................................. 27
5. Conclusion ................................................................................................................................... 29

Bibliography ....................................................................................................................................... 30
1. Introduction

“Over himself, over his own body and mind, the individual is sovereign”. (J.S. Mill)

This thesis looks at the methods by which governments and the private sector do and can influence individual decision makers. It explores a number of critiques in addition to the justifications offered by proponents of such techniques. As such the subject of paternalism is often associated with the technique of nudging. This is largely due to the motivations behind influences; paternalism can be described as “the interference of a state or an individual with another person, against their will, and defended or motivated by a claim that the person interfered with will be better off or protected from harm” (Dworkin 2017). Since many governments initiated behavioural steering techniques are carried out in the name of ‘societal good’ they might be classified as paternalistic measures (The Behavioural Insights Team 2018). The argumentation in this thesis however, is distinct from traditional debates on paternalism since it deals not only with interferences deemed to be in the interests of those interfered with, but with influences perpetrated for all manner of reasons.

In this thesis, I deal not only with behavioural influences carried out for a number of reasons but done so in a number of ways. Chapter 2 offers a background in behavioural economics; I explain how psychology and behavioural insights might be employed as a means of exploiting individual cognition, therefore shaping behaviour. Chapter 3 provides some conceptual clarity; differentiating techniques of nudging, boosting and market advertising. Chapter 3 also highlights the difficulty of defining such techniques. Existing definitions, of nudging especially, often conflict with one another or do not accurately describe nudges in practice. As a result, ethical justification can become inapplicable. Chapter 3 explores these ranging issues, before in chapter 4 I attempt to address them. Chapter 4 provides arguments against opaque methods of behavioural steering on grounds of republican freedom and individual autonomy. I do not disagree that individuals may indeed be influenced in ways that benefit their own wellbeing however, individuals must consciously be aware of influences in order to make free choices. Wellbeing, I argue, is in part defined by freedom; as such influences that promote wellbeing must therefore be transparent.

Perhaps the most important implication, discussed within this thesis, is that sub-conscious or manipulative methods of behavioural influencing can potentially have long term effects on individual autonomy. Where choices are continually removed or influenced, individual decision makers are reduced to “the instruments of someone else’s will” (Grill 2014). Effectively, humans become sheep, herded by the government towards idealised models of society, or by the private sector, propped up by ‘rationalized’ consumers (Bradbury et al. 2012; Ball 2001; Walkerdine 2003).
These potential outcomes provide further weight to my conclusion that transparency is morally essential.

The discussion of whether psychological behavioural influencing techniques can be ethically permitted is not only theoretically interesting but highly relevant. The British government created a Behavioural Insight Team (BIT) in 2010, that has since become a ‘social purpose company’. The BIT is responsible for many innovative programmes that draw on psychological behavioural insights with huge effect. Often these programmes or policies have international implications. The BIT’s *International development and behavioural insights, Summary report 2017-19* describes their recent work establishing government behavioural insights units in Indonesia, Guatemala and Bangladesh, furthermore, it describes its efforts to ‘combat corruption in Nigeria’ (The Behavioural Insights Team 2019). Clearly, behavioural insights, used as tools for influencing changes in behaviour, provide a highly relevant topic for ethical consideration. Before private organisations and governments feel entirely secure in their use of behavioural insights, it is essential that they be evaluated from a normative ethical perspective.

2. A brief background in behavioural insights and choice steering

The school of behavioural economics is a subcategory of psychology that attempts to understand the various social and situational influences behind a person’s decision making i.e. to gain ‘behavioural insight’. In strategic behavioural influences, such as nudges,

“the focus shifts from employing behavioural insights as a prognosticator of human behavioural response to being the key in the practice of shaping that behaviour”

(Kosters and Van der Heijden 2015, p.278).

Knowledge of behavioural insights can be used to devise methods of behaviour steering/influencing/shaping.

Research and the implementation of behavioural insights, as a steering technique, has found that the laying out of choices can have a significant impact on outcomes (Haugh 2017). For example, studies have shown that using ‘default setting’ will significantly influence results. Johnson & Goldstein’s 2003 study demonstrates this in practice: In many European countries organ donation is the default setting (individuals have to ‘opt out’ rather than ‘opt in) as a result there are far greater numbers of registered organ donors in these countries than in others (Johnson & Goldstein cited in Kosters and Van der Heijden 2015).
By offering the same options but in terms of ‘opt out’ rather than ‘opt in’, governments or businesses can quite easily influence certain decisions. Proponents of nudge often argue that this type of influencing does not remove an individual’s free choice (Thaler & Sunstein 2009). Research in behavioural economics explains the psychological reasoning behind phenomena such as default setting. Thaler and Sunstein explain that often “many people will take whatever option requires the least effort” (ibid, p.85). While the psychological underpinnings of behavioural economics may at times seem obvious, it often requires specialist knowledge and insights into human behaviour to calculate the ways in which choice options are to be constructed. Thaler and Sunstein describe the constructors or designers of choices as ‘choice architects’ (ibid). Many factors can influence a person’s decisions however; considerable research shows that the way in which choices are presented (choice architecture) can have a significant effect (Gilovich, Dale and Kahneman 2002).

In this thesis I look specifically at choices influenced by choice architecture, rather than by other means i.e. bribery, coercion, legislation, etc. Many organisations, governments and businesses alike, have drawn on behavioural insights and research in choice architecture to steer people’s decisions making. Nudging, boosting or just plain market advertising are examples of where behavioural insight is utilised as a means to steer decision making (Thaler and Sunstein 2009; Grune-Yanoff 2017; Rucker et al. 2007).

It should now be clear what is meant by behavioural insights and behavioural economics. In the following sections I shall provide a basic description of some of the core beliefs about human decision making, which inform the construction of choice architecture. This is important when it comes to defining and differentiating between behaviour influencing techniques; as they stem from research programs that differ slightly in their interpretation of heuristics and cognitive processing (Hertwig & Grüne-Yanoff 2017).

2.1 Two systems

The first idea that often reinforces behavioural steering is the belief that humans use two systems of cognitive processing (Thaler and Sunstein 2009; Soloman, 2002). Of these two systems; system 1 works quickly, on instinct while the other (system 2) takes longer to rationally puzzle and solve more complex problems. The quick system is described by Thaler and Sunstein as the Automatic system whereas the slower one is called the Reflective system (2009, p.17-22). Both systems are important: some situations require a swift response (i.e. catching a ball), while others need to be considered more carefully (i.e. solving a difficult maths problem).
Humans have the ability to instinctively employ one system or the other, however, they can make mistakes. To clarify, consider the example: sometimes people ‘think too much’ before taking a difficult shot in sports i.e. badminton, they miss because they were too busy analysing the shot (using system 2) rather than reacting (using system 1). A lot of physical activities do not require system 2- reflective thought; once a person has learned the rules of a game they can, through practice and repetition, gain skills which become ingrained in their automatic system (system 1). Once these skills or techniques, and knowledge of where to apply them, has become embedded in system 1, system 2 is no longer required for such activities.

Often people rely on their automatic systems. If we subscribe (like Thaler and Sunstein) to the belief that people use a ‘dual-system’ of cognitive processing, we must understand that the automatic system takes over much of our behavioural responses: it does so largely because of its speed and because it takes less effort. Gordon Foxall writes:

‘On the understanding that individuals minimise effortful activity, systematic processing [the processing done by the reflective system] is likely only when the person is highly motivated and has the cognitive capacity and resources to engage in it’. (1997, p.41)

Because people avoid ‘effortful activity’ and do not always possess the ability to understand some choices, they may wrongly utilise their automatic system. Knowledge of this dual-system processing can be (and often is) exploited by choice architects. This can be seen in many cases: consider long and complicated forms; people may skim read or simply skip to the bottom where they sign. The idea of ‘reading the small print’ comes from this commonly used technique; that exploits people’s avoidance of effortful systematic processing or their inability to adequately understand; when they fail to read all of the stipulations on a contract or product.

2.2 Humans and Homo economicus

The understanding of two systems helps us to understand a second concept- people are not always rational. Here, since rationality is a central concept in decision making as understood by behavioural economics, it requires some form of explanation (there exists differing conception between and among academic fields). The Rationality Principle is often said to be “the fundamental principle of economics” however, a succinct definition of this principle is more difficult to come by (Lagueux 1997). Generally speaking, a person is said to behave rationally if their actions further their attempts at chosen goals. Lagueux describes rationality as a ‘propensity’ or ‘capacity’ to maximise economic gain (1997). J.S. Mill describes an economic man as a being who desires wealth; he will be capable of judging efficient ways of acquiring wealth and therefore pursue means
of achieving his goal (cited by ibid). The ability to accurately judge means of reaching said goals (often financial when discussed in economics) can be understood as a rough definition of rationality. While people can make rational decisions, both automatically and reflectively, they will often misjudge situations and make mistakes- they are after all only human. Thaler and Sunstein point out that while most people are aware of human fallibility, “many people seem at least implicitly committed to the idea of homo economicus” (2009 p.6). Thomas Pogge offers a succinct definition of this idea, often prominent in economics: “an individual who, single-mindedly and rationally, seeks optimally to satisfy his preferences (2004, p.29)”. He adds “Such imaginary creatures are not good approximations of persons in the real world” (ibid). Thaler and Sunstein agree; in practice a number of elements influence our decision making- not purely rational self-interest (2009). Sadly, it is well beyond the scope of this paper to fully explore the reasons for irrational decision making, suffice to say they are numerous and most academics agree that all humans may on occasion act irrationally.

Knowing that humans can lack ‘rational judgement’, choice architects face a choice. 1. They can choose to design choice architecture that will steer individual decisions towards rational choices i.e. toward options that with direct individual’s towards achieving their own goals. 2. They could remain neutral by randomising design or attempting to leave individual decisions uninfluenced. 3. They could further their own ends by exploiting individual weaknesses by “harness[ing] system 1’s deficiencies” (Grüne-Yanoff and Hertwig 2017, p.8).

Returning to the case of organ donation, it can be claimed that the governments that employed choice architecture, in the form of default setting, did so to achieve rational outcomes, though perhaps not for decision makers individually but on a societal level. The classical economics definition of rationality (see footnote 1), may well work in the field of economics however, rational decisions might not always further individuals but society as a whole when understood in terms of behavioural economics. Likewise, the ‘rational’ goals of society may not be motivated on a purely financial basis but on grounds of perpetuating a functional or healthy society. I think that, by making a distinction between humans and homo economicus, Thaler and Sunstein begin to generate a broader understanding of rationality which seems necessary when discussing policies informed by behavioural economics. An alternative understanding of rational decision making can be defined in terms of satisfaction of needs rather that fulfilment of goals. Menger explains; “it is the endeavour to ensure the fullest possible satisfaction of their needs” (cited by Lagueux 1997, p.3). This account of rational decision making appears to appreciate that not all needs can be satisfied financially, so rationality might be the capacity to undertake activities or make choices that will progress an individual towards satisfaction of their needs in the best possible way. Satisfying needs can be seen
similarly to enhancing wellbeing. Regardless of how rational choice is understood, what seems clear is that choice architects are often in a position of power, whereby they can influence individual decision makers in whatever way they think best.

2.3 Attitude and decision making

_A veritable bounty of research has demonstrated that one’s attitude influences one’s behavioural intentions, and that behavioural intentions best predict behaviour._ (Rucker et al. 2007, p.74)

Rucker et al. believe that advertising can be treated as a science that draws on literature from psychology. This is rather similar to Thaler and Sunstein’s belief that behavioural economics, can inform choice architecture; influencing a person’s decision making. Rucker et al. emphasise the importance of a person’s attitude in informing their behaviour (which includes decision making). Like Thaler and Sunstein they describe the dual-system of processing; focusing on the capacity to be persuaded. They describe a central route to persuasion; like system 2 it “involves effortful scrutiny of issue-relevant information” and a peripheral route that is similar to system 1, relying on “mere exposure (Zajonc, 1968) and use of decision heuristics (Chaiken, 1980)” (2007, p.75-77).

Rucker et al. explain how the attitudes formed in a more systematic manner will last longer and are more resistant to criticism.

‘Attitudes formed through the central route, relative to the peripheral route, are more inclined to persist over time, stand resistant to attempts to change them, come to mind quickly, be included in consideration set prior to choice, and be more likely to influence and predict other judgement and behaviours (for a review see Petty et al. 1995)’ (ibid, p.79)

Advertisers may, for this reason, attempt to persuade consumers through the central route. This method of persuasion may be more difficult, as already stated people need to be highly motivated to engage in this type of persuasion however, its effect on consumer attitude is much stronger. This kind of persuasion can make consumers feel better about their choices as they have a more informed and systematic cognitive base for their consumer choices; which can be defended if challenged (ibid, p.79-80).

It is important to be aware of the attitude strength of a decision maker, either before attempting to influence them or after having done so. Attempts to influence or strengthen system 2 are uncommon. Grüne-Yanoff and Hertwig explain why: Firstly, system two attitudes are strong
(2017). As Rucker et al. explain, those formed through the central route are difficult to influence or change (2007). For this reason “debiasing attempts are often seen as futile” (Grüne-Yanoff and Hertwig 2017, p.8). The second reason cited, is that due to its difficulty, system 2 interventions can be costly or time consuming. Nudges, which alternatively ‘harness system 1’s deficiencies’, have a much larger scope and can be implemented cheaply (ibid). This may be a ‘selling point’ for nudging however, the changes to a person’s attitude, made by influencing system 1, may be short lived and superficial, requiring constant ‘top ups’. Boosting attempts to influence system 2, influencing on a deeper and systematic level. Market advertising may attempt either type of influence, largely this is dependent on the product or service being sold and to whom.

2.4 Heuristics or Rules of thumb

Heuristics and biases form a “core element of behavioural economics” (behaviouraleconomics.com). Rucker et al. provide a basic definition: “Heuristics represent mental shortcuts that allow for simple decisions based on a rule of thumb” (2007, p77). Rules of thumb are general beliefs, often gained through experience, which can guide people’s choices i.e. “Experts are usually right” or “higher prices mean better quality” (ibid). These heuristics arise from the interplay between the automatic and reflective systems (Thaler and Sunstein, 2009). Initial experiences may well support the generated rules of thumb however, the use of resulting heuristics, to inform future choice, may not be rational or reflective. Indeed people may form biases that colour their future decisions without full and careful consideration of current evidence and events. These biases are known as cognitive biases, they often skew human judgement, resulting in irrational behaviour. Behavioural steering techniques often rely on knowledge of common biases in order to influence choices before they are made; it is therefore existing biases, rather that irrational behaviour that provides an opening for choice architects.

3. What differences in ethical challenges do nudges, boosts and market advertising face?

By now some of the mechanics of what I have grouped and labelled (psychological) strategic behavioural influences should be clear. The purpose of this chapter is to define each technique separately and then to assess whether ethical challenges may be broadly applicable. While proponents of boosting strongly emphasize the differences between it and nudges, they concede that “educative nudges and short-term boosts largely overlap” (Grüne-Yanoff and Hertwig 2017, p.5). Likewise, Whitehead et al. explicitly claim, “Nudge promotes the use of well-rehearsed techniques of corporate advertising as bases for promoting non-compulsive influence” (2012, p.303). It is
because of this overlap, that I address the selected techniques simultaneously in this thesis however, it makes defining them particularly challenging. In light of this semantic challenge, I have elected to draw out the distinctions through comparisons. I therefore begin by offering a general definition of nudging, which I subsequently compare to boosting. This comparison includes an assessment of whether boosts truly manage to circumvent the normative implication commonly associated with nudging. In the following section, I compare nudges and market advertising. Again, I consider whether advertising is subject to the same critiques as nudging or indeed whether it brings about further ethical implications.

3.1 What is a nudge?

A common place to begin, when understanding nudge, is with the definition offered by Thaler and Sunstein:

“A nudge, as we will use the term, is any aspect of the choice architecture that alters people’s behavior in a predictable way without forbidding any options or significantly changing their economic incentives. To count as a mere nudge, the intervention must be cheap and easy to avoid. Nudges are not mandates. Putting fruit at eye level counts as a nudge. Banning junk food does not.” (Thaler and Sunstein, 2009, p.6)

This definition has come under heavy scrutiny for a number of reasons; it is not technical, it does not always align with the examples offered in *Nudge* and it is ‘fuzzy’ (Saghai 2013a, Kosters and van Heijden 2017; Selinger and Whyte 2011). Regardless, this definition captures the essences of nudging underpinned by what Thaler and Sunstein call *libertarian paternalism*.

3.1.1 Libertarian Paternalism

Libertarian paternalism attempts to offer a ‘third way’; one that incorporates libertarian principles of freedom of choice while also steering the behaviour of people in the direction of what might objectively be deemed ‘their own best interests’. As Thaler and Sunstein understand it ‘rational’ choices appear to coincide with individuals best interests. They offer numerous examples of people who lack the ability to make choices that would objectively be considered in their own best interests, i.e. overeaters, drinkers, smokers, etc. Nudges that, for example place healthier food at eye level, help to steer individuals’ choices in a healthy and therefore rational way, without restricting their freedom to select less healthy food (2009).
3.1.2 ‘Fuzzy’ or ‘mistaken’ nudges

Though there certainly is a clear concept of what a nudge is, there is some confusion when people design and talk about nudges in practice (Selinger and Whyte 2011, p. 924).

As Selinger and Whyte point out, the problem with defining nudges often hinges on how they are implemented. Many policies, described as nudges, do not actually adhere to libertarian paternalistic ideals. Taken purely from various examples offered in Nudge, we can see that interventions such as Dollar a Day and stickk.com offer financial incentives. Dollar a day offers teenage girls one dollar for every day that they are not pregnant thus deterring teenage pregnancy. Stickk.com helps individuals to stick to their self-specified goals i.e. losing weight, by holding an allotted amount of the individual’s money which will be returned if the goal is met but sent elsewhere if not. A further example, of special licenses for motorcycle users who do not want to wear helmets, would require choice-restricting mandates (2009). Selinger and Whyte argue that is important to know what exactly constitutes a nudge, this is so that citizens can “consider critically whether they should support ‘alleged’ nudge policies” (2011, p. 924).

I agree with Seliger and Whyte on this point however, I do not agree with their definition. They believe nudges to be: inexpensive, easy to opt out of, void of financial incentives, transparent and “designed to help people live according to their best interests” (2011, p. 926). This understanding of nudge is rather idealised since many nudge policies offer small financial incentives, likewise the difficulty of opting out may vary from person to person; this makes it a blurry scale on which to judge what counts as a nudge. Perhaps most importantly, many nudges in practice are not transparent. Even Thaler and Sunstein’s example of Lake Shore Drive, cited by Seliger and Whyte as an authentic nudge, does not steer individuals in a transparent manner (2009; 2011). Road markings on Lake Shore Drive are designed to slow down travelling cars, they do so sub-consciously by having evenly spaced lines on the road that then get closer in proximity nearer to the bend. The closing gaps between the lines gives drivers the illusion that they are travelling faster than they are, they therefore slow down. This example demonstrates how nudges regularly fail to meet the requirements of transparency.
3.13 A technical definition

Since nudges in practice rarely meet the criteria of theoretical and idealised nudges Yashar Saghai reformulates a technical definition of nudging:

“Nudge. A nudges B when A makes it more likely that B will ϕ, primarily by triggering B’s shallow cognitive processes, while A’s influence preserves B’s choice-set and is substantially non-controlling (i.e., preserves B’s freedom of choice).” (2013, p.491)

This definition is focused especially on the element of non-control. Saghai explicates Thaler and Sunstein’s loose phrase ‘easy to avoid’ in terms of substantial non-control. Here he focuses of how individuals should retain a full set of options while also being free from coercion, compulsion, incentive and disincentives (financial or otherwise) (2013a).

3.1.4. A loose definition of nudging

Given the difficulties of defining nudges in practice I, like Thaler and Sunstein, offer a rather loose definition that I hope casts rather a wide net. I do this primarily so that we can discuss nudges in practice without denying criticisms on semantic grounds, especially since nudges that have actually been implemented rarely fit within tightly defined frameworks. Furthermore, I define nudges in general terms so that we can understand them as both ethically viable and nonviable. I believe, what makes behavioural influences/ nudges permissible or impermissible is not bound by definitions, but rather by the implementation, consequence and its situational qualities. A parallel example can be drawn in cases of murder, defined as the intentional killing of another. While never good, situational qualities can make it ethically permissible i.e. whether it was done in self-defence.

Nudging, as I understand it, occurs when an individual’s choices are intentionally influenced by another party in ways that have been strategically implemented, drawing upon the field of behavioural science. The third party could be governmental, corporate or individual and their intentions may be either to benefit society, the individual in question or themselves. Nudges are not
bribes, and should not provide significant financial incentives, nor should they be coercive or option restricting (options should not be removed). Nudges should not contain explicitly false information (lies), though they may withhold certain information or present facts in favourable ways. Nudges are aimed at altering one specific behaviour, rather than developing individual’s capacities or changing their values.

3.2 How do nudges differ from boosts?

Grüne-Yanoff and Hertwig are largely responsible for the concept of boosting and as such they are clear in drawing distinctions between boosts and nudges. In this section I shall review the seven dimensions of boosts, offered by Grüne-Yanoff and Hertwig, which make it conceptually distinct. Furthermore, I shall consider the ethical implications for nudges as compared to boosts.

3.2.1 Intervention target

Perhaps the most important feature of each intervention (nudge or boost) is its intention or target. Nudges intend to alter people’s behaviour when it comes to making decisions about improving their ‘health, wealth or happiness’ (Thaler and Sunstein 2009). Alternatively, the target of boosts is individual’s capacities. Boosts work by providing individuals with the necessary mental or physical tools/ capacities to make good choices; these tools might take the form of financial literacy education, to better understand complicated household bills or physically providing information in the form of leaflets. Boosts exist as influences since, without them, individuals could make very different choices.

The paternalistic element of nudging i.e. altering individual’s behaviour by steering choices towards decisions deemed to be in their best interests, has received much criticism. Largely, this is due to the fact that altering a person’s behaviour violates their autonomy. Autonomy can be understood in a number of ways, one of which is ‘freedom of choice’ (Vugts et al. 2018). Nudging, it can be claimed, reduces freedom of choice. In nudges such as Dollar a Day, financial incentives are used, combined with social pressure it might be seen as difficult to avoid this nudge. This violates Thaler and Sunstein’s principles of libertarian paternalism, furthermore, Saghai attempts to ‘Salvage the concept of nudge’ by excluding interventions that have a stronger degree of control than ‘substantial non-control’ (2013a). Regardless of how we define nudges, many nudges in practice appear to exert more than substantial non-control, they therefore reduce individual’s freedom of choice.

While reducing freedom of choice may be justified on consequentialist grounds i.e. the outcome was good and therefore the means were just, it seems clear that autonomy can be restricted by nudges. In fact, the very intention of a nudge is to target behaviours and decisions that would lead to
different outcomes if un-steered. Boosts however, do not have the same intervention target and, for this reason, are not subject to the same critiques. Boosts often provide more information or foster critical capacities for good decision making; this means individuals retain full freedom of choice as boosts do not exert any means of control. Individuals may freely choose whether to engage in boosts but without their engagement boosts have no effect. If individuals do engage with a boost, they still retain freedom of choice when it comes to deciding what action to take.

3.2.2 Roots in research programs and evidence

Nudging has its roots in the heuristics-and-biases research program which gives a rather condemning view of human decision makers, capable of irrational choices and easily susceptible to biases (Thaler and Sunstein 2009; Grüne-Yanoff and Hertwig 2017). Alternatively, boosting is informed by the simple-heuristics research program, which demonstrates the efficiency and efficacy of human dual processing. While humans have the potential to behave irrationally, they often make difficult and pressured choice with the help of simple heuristics. Fire fighters and nurses, for example, rely on rules of thumb or heuristics in situations where there isn’t time for reflective thought. This is often essential and usually yields positive outcomes (Klein cited Grüne-Yanoff and Hertwig 2017). These programs do not directly contradict one another since it is clear that individuals are both capable of rational and irrational decision making however, these research-approaches examine human decision making from opposing viewpoints. Paul Hammilton cites this as a potential issue however, he sensibly points out that we should “let the psychologists decide which policy is better grounded in cognitive science” and instead treat boosts and nudges as distinct actions, each carrying “distinct moral implications” (2018). This is exactly what I shall do.

Regardless of which research program is best psychologically supported, there may be an ethical danger in viewing individual decision makers collectively as irrational or incapable of overcoming biases. Luc Bovens describes the danger of ‘infantilisation’ (2009). He does so in terms of long-term effect; treating individuals as creatures whom often make irrational choices may lead to nudgers overstepping individual agency. This rendering human beings as incapable of rational decision making based on fact or evidence, Bovens likens to making children out of adults. Acting with this notion in mind may lead to choice architects believing that only through their intervention is society capable of positive progression.

Vugts et al. demonstrate how autonomy might also be thought of as agency. They explain, “An agent is- at least to some extent- in control and capable to navigate in the world on a voluntary basis” (2018, p.9). This understanding potentially provides and obstacle to the heuristics-and-biases research program which “show[s] decision maker[s] as systematically imperfect and subject to
cognitive and motivational deficiencies” (Grüne-Yanoff and Hertwig 2017, p.2). In the case of nudges, individual agency is often overruled when choice architects view decision makers as irrational subjects and make ‘better choices’ for them.

A further issue is that cognitive science and behavioural economics suggests people are not capable of always making rational and informed choices, for this reason choice architects may view their interventions as benevolent, since they intend to bring about rational outcomes (Hammilton 2018). Unfortunately, good, well thought out, rational choice for one person may not be same as those of another. Choice architects have to project their conceptions of what is ‘good’; the problem being that ‘conceptions of the good’ depend on ‘subjective assessment’ (Kosters and Van Heijden 2015). In many nudges, individuals are viewed collectively and denied to opportunity to navigate choices in the way they best see fit. For this reason, their agency is undeniably curtailed.

Boosts operate on a more individualistic level since they are offered and can be rejected. If individuals choose to engage with a boost, i.e. by reading the simplified information on the back of an electricity bill, they retain freedom of choice when deciding how they might reduce their monthly bills. They might choose to do nothing, furthermore, they might realise that they have a good deal with their energy provider. Either way, the boost is directed at the individual and actually enhances their agency by providing additional capacities to support their ‘navigation in the world’.

So, in summation, Boosting is essentially an attempt to provide an individual with the relevant information, (facts and evidence), with which to make an informed choice. The individual is left free to ignore the evidence, to incorporate it into their decision making, or to reach their own opposing opinions based on it.

3.2.3. Causal pathways

Grüne-Yanoff and Hertwig use the term causal pathways to refer to the way in which a person is influenced. Nudging alters the exogenous properties of choice architecture i.e. things outside of the decision maker. This might be: the ordering of lists; default settings; the phrasing of statements. In contrast, boosts target human cognition, the environment or both (2017).

When educational boosts offer information or provide skills for individuals to make better decisions, they alter a person’s cognition. This type of influencing could be subject to criticisms concerning indoctrination as a threat to autonomy as self-constitution (Vugts et al. 2018). Self-constitution refers to autonomy in terms of individual identity, authorship and personal values. Baldwin has shown concern, believing, “when nudging goes beyond the supply of simple
information, or reminders… what is involved is the nudger acting in a semi-covert or covert manner to further the nudger’s own conception of the target’s welfare- which may involve re-shaping the target’s idea of their own welfare” (2014, p.846). This is a legitimate concern in the case of nudging, as previously pointed out, choice architects have the ability to project their own conceptions of the good onto individual targets. In the case of boosting however, interventions do not extend beyond ‘simple information, or reminders’, for this reason they do not carry the potential for indoctrination in the same way as nudges. However, that being said, by providing misleading or partial information boosts are just as capable of being used for indoctrination. Nothing, no system or technique, is entirely impervious to corruption. This is an important point to keep in mind.

3.2.4 Normative implications

What Grüne-Yanoff and Hertwig refer to as ‘normative implications’, is the requirement for transparency. They list this as the final dimension of what make boosting distinct form nudging however, I think it is important to address it now since the normative issues surrounding behavioural steering discussed so far are highly dependent on a subject’s awareness of said steering influences. Grüne-Yanoff and Hertwig stipulate that boosts require transparency to be considered as such. They claim that in order to engage with a boost a person must first be aware of it and its intended outcomes. This differs to nudging, which may often work on a sub-conscious level (Goodwin 2012).

While Thaler and Sunstein call for transparency (2009; Thaler 2015), many of the examples used in Nudge do not fully meet this criterion i.e. Lake Shore Drive. The non-transparent element of this nudge is the illusion of greater speed. This is what makes it effective i.e. makes drivers slow down. Initially, it is difficult to see why this would be ethically problematic, as slowing down appears to be in everyone’s best interests. The issue however is that it operates on a sub-conscious level, drivers therefore do not learn when to slow down, nor do they give it any reflective thought. More worryingly, were road-marking-nudges to be used frequently, driver’s perceptions of speed could be hampered or warped, leading to reduced capacities for driving and potentially more accidents. This lack of development is dangerous in the physical sense, but it also carries dangerous implications for moral development (Bovens 2009; 2012). Drivers affected by opaque or manipulative nudges may rely heavily of nudge interventions, undermining the need to develop moral considerations for other drivers, passenger, pedestrians or themselves.
3.2.5 Assumptions about cognitive architecture

Nudges, due to their roots in the heuristics-and-biases research program, make a number of assumptions regarding the ‘cognitive motivational deficiencies’ of individual decision makers i.e. that they are susceptible to certain biases. Based on these assumptions, choice architecture is created with the intention of steering certain individuals i.e. those for whom the assumptions are correct. For this reason, nudges are limited in their scope since they are only effective in steering certain (biased) individuals. Boosts do not make the same assumptions, nor do they use them as a lynch pin, around which individuals can be manoeuvred, instead they target specific capacities. By targeting capacities, boosts are specific to any individual who lacks said capacities and wishes to develop them. Boosting operates under the encouraging belief that individual decision makers have malleable cognitive architecture, therefore they are capable of changes in behaviour and thought. These underpinning beliefs do not have direct ethical concerns however as in the case of roots in research the underlying assumptions about individual’s capacities may inform the type and level of intervention exercised by the choice architects.

3.2.6 Empirical distinction criterion (reversibility)

Grüne-Yanoff and Hertwig explain how boosts and nudges could be employed to achieve similar ends i.e. changes in behaviour. They offer an example of nudging, Save More Tomorrow (SMT) intended to increase the number of people saving for their retirement (Thaler and Benartzi cited by Grüne-Yanoff and Hertwig 2017). It operates on a number of assumed ‘cognitive motivational deficiencies’ such as the present bias- the stronger preference people have for receiving rewards now rather than in the future. Knowledge of these biases is utilized in SMT to nudge individuals towards saving however, it does this only once. Alternatively, boosting methods that could also be employed as a means of encouraging individuals to save for their retirement, might inform them in other decisions related to savings. The capacities developed by boosting interventions can be transferred to other choices and therefore are not so easily reversible. Grüne-Yanoff and Hertwig state that, “competences can be specific to a single domain” or “generalise[d] across domains” (2017, p.5). This certainly seems plausible if not likely, however, they explain that there has been little empirical research to support their hope that “the desired behaviour may ‘survive’ the removal of the scaffolding choice architecture” (2017, p.7). What is more certain is that interventions that develop capacities rather than exploiting biases promote individual agency through transparent means.
3.2.7 Programmatic ambitions

This final dimension of boosting, which makes it distinct from nudging, is very much linked to the reversibility criterion. It refers to the fact that nudges are deployed with the intention of influencing decisions in one specific context i.e. the choice of whether to commit to a pension plan. Employers in the UK are legally required to automatically enrol all employees into a pension savings scheme, this is a nudge since all employees must also be legally informed and have the opportunity to ‘easily’ opt out. This nudge has been hugely effective in increasing the number of retirement savers from 49% to 86% (Service 2015) however, it has been effective only in increasing the number of retirement savers. Nudges might have a mass impact, but they can only influence in one context. Alternatively, boosts can affect individual’s capacities to make decisions in generalizable contexts, i.e. in all savings related situations.

The transferability of capacities is no doubt a virtue of boosting however, I think Grüne-Yanoff and Hertwig downplay the potential impact that nudges might have on future decisions. Joel Feinberg argues that freedom of choice is undermined if or when ‘fecund’ options are removed (cited in Saghai 2013b). Fecund options are those that affect other choices i.e. whether to study for an important exam. If a student does not study and fails as a result they may not be accepted into the university of their choice, thus their enthusiasm for academia is undermined; subsequently restricting future career options. As this example demonstrates, nudges might only influence one choice directly however they can produce a ripple effect. Boosts too have a ripple effect however, the initial act in boosting can be said to have been made with greater freedom of choice. In certain covert nudges, individuals may not even be aware of decisions that they have made and how they may impact on future choices. The ethical implications for nudges can become rather expansive if we accept Feinberg’s argument.

3.3 Market advertising and nudge

A nudge is a specific way of choice architecture [...] that gently pushes the behavior to a predictable point, without omitting other choices or influencing economic motives [...] Based on the aforementioned, we could see the applicability of nudges in advertising. Theoretically, it is the best way to advertise, maybe the essence of advertising itself.” (Christos Panousis 2016)

Like Panousis, I can see how nudging and advertising are incontrovertibly linked; this is not to say though that all adverts are nudges. Below I aim to explain the key differences, as I see them, between nudging and advertising. Additionally, I raise some of the ethical issues that I find most
troubling with the implementation of behavioural steering in market advertising. A final point for consideration is that adverts and nudges, while distinct, can at times overlap. This I demonstrate before progressing to chapter 4.

3.3.1 Purpose

The purpose of advertising differs from that of nudging and boosting. Freitas writes, “advertising is one of the elements of the communication strategy of a brand” (2008). She goes on to state the four hierarchical stages in the communication process, as described by the DAGMAR model: 1. Bring about consumer awareness of the product, 2. Develop the understanding or comprehension of a consumer in regards to the product, 3. Develop consumer conviction and positive disposition towards the product, 4. Ensure the consumer acquires the product (Belch and Belch cited by Freitas 2008). The panicle goal of advertising is to ensure a product is sold; the sale of products is what makes a brand successful. If we accept this explanation of advertising it seems clear that its purpose in distinct from that of nudging which seeks to alter specific behaviours that are usually believed to be in the best interests of the individual; advertising seeks to alter consumer behaviour to that which is in the best interests of the brand.

Nudges could be seen as relatively mild influences when compared to the DAGMAR model described above. Nudges target one specific behaviour and as Grüne-Yanoff and Hertwig point out, the new behaviour rarely survives when choice architecture is removed (2017). In comparison, advertising intends to develop a “conviction, where positive disposition towards the product is developed” (ibid, p.43). This kind of conviction can be seen in consumer brand loyalty however, it raises ethical questions in regard to autonomy.

Saghai describes an example from *The Nudge blog*, which shares many of the elements used in advertising to create brand loyalty. This example is called *Asparagus-Lovers*. It describes how it is ‘suggested’ to participants of a study that they enjoyed eating asparagus the first time they tried it, during their childhood. This suggestion ‘primes’ the participants, creating a false memory. Consequently, participants report greater or increased liking for asparagus, not to mention a willingness to pay more for it (Saghai 2013a). Advertising taps into many areas of psychology and is therefore practiced when it comes to priming consumers. This kind of influencing is perhaps more effective, but it also exerts a greater level of control over individuals. Sagahi rejects Asparagus Lovers as a valid nudge, instead he believes that it is “an example of an arranged hostile environment” (ibid. p.491). A hostile environment for heuristics is one where clues are misleading or lacking (Stanovich cited by ibid). Adverts rarely present all of the facts when describing a product and they employ a range of behavioural insights to ensure heuristics can be exploited, for
this reason they do not offer the same level of non-control as nudges and instead might be classified as behavioural ‘prods’. Saghai claims that, “Marketers often prefer prodding to nudging”, which suggests that behavioural influencing techniques in marketing are less concerned with individual autonomy (2013, p.491).

3.3.2 Accountability

Accountability is an important feature in differentiating between nudges and advertising. Before addressing this conceptual distinction, it is first worth noting that nudges can be separated into two distinct types, each subject to slightly different accountability (Kosters and Van der Heijden 2015). Type one nudges are those that the governments implements so as to steer the choices of individuals towards outcomes that are in their own best interests. Type two nudges are government interventions that steer individuals’ behaviour towards collective ends that are in society’s best interests. I would like to add a third and fourth type of nudging to those suggested by Kosters and Van der Heijden. These would be nudges implemented by private organisations; either for the individual’s best interests or for collective ends. A possible fifth type of nudging would be the use of behavioural influencing techniques that steers individuals towards behaviour that is in the interest of the choice architect. This fifth type no longer meets the libertarian paternalistic goals however the Nudge blog defines them still as nudges:

“While nudges cited in the book [Nudge] are intended to do exactly this, [“…help people make decisions that make them better off as defined or judged by themselves…”] nudging takes place in [a] variety of realms where the nudger’s explicit goal is to promote [the nudger’s] own welfare” (cited by Saghai 2013a, p.488).

Thaler and Sunstein address the issue of self-serving nudgers within Nudge (2009); they do so in a section entitled Evil Nudgers and Bad Nudges. Thaler and Sunstein do not endorse Bad nudges however their response is rather weak. They suggestion is: “try to line up incentives when we can, and employ monitoring and transparency when we can’t” (2009, p.242). By ‘lining up incentives’, Thaler and Sunstein are suggesting that the bad nudges can be legitimised, if they also happen to benefit individuals. Thaler and Sunstein also point out that it would not be in the best interests of organisations or businesses to nudge individuals in damaging directions. This defence is typically libertarian, and indeed Thaler and Sunstein argue; “if private-sector interests are just following the invisible hand in furthering the interests of their customers, what’s the problem?” (p.243). The problem is that incentives do not always line up, as in cases where products are overly expensive, dangerous or create unnecessary want. Thaler and Sunstein suggest that in such situations ‘monitoring and transparency’ are required (ibid). The problem with this solution is that
nudges are not always transparent. Furthermore, monitoring nudges might sound fairly straightforward but countless businesses employ nudging techniques, not to mention the difficulties involved in governmental monitoring. To monitor all nudges would be impossible; the closest realistic option might be to legislate the use of certain nudges. By placing legislation on nudging, libertarian principles give way to paternalism.

Due to the ‘covert nature’ of many nudges, accountability becomes very difficult. Kosters and Van der Heijden note that the BIT’s director Halpren ‘acknowledges this weakness by arguing that ‘tacit acceptance’ is required from the public’ (2015, p.281). This kind of acceptance might not be so unreasonable when we consider the nature of democratic governments, since the public, to some extent, has to trust in government officials, acting on their behalf. This level of trust, however, cannot be expected of private-sector nudges or in cases of market advertising.

Market advertising, which may draw on behavioural insight, does not operate with the purpose of enhancing or empowering rational decisions, for this reason it cannot receive the same level of public trust that might be shown to government nudges. Instead, businesses and organisations that implement influencing strategies are held accountable by law. Within the EU, advertisers are constrained by a number of laws, some of which are specific to the county in which they operate (European Union Country Commercial Guide 2018). The European Commission has adopted a number of directives that attempt “to establish minimum objective criteria regarding truth in advertising”, furthermore, several regulations limit advertising for junk food aimed at children, prohibit tobacco commercials and regulate nutrition and health claims (ibid.). In light of this, the different modes of accountability can be cited as a distinguishing feature of nudges and adverts.

3.3.3 Roots in research and methods of implementation

To some degree market advertising has greater ‘creative’ freedom than nudging, this is because it does not only draw on the heuristics and biases research program, but on a range of research in advertising, persuasion, attitude change, attitude strength and theories of emotions. This helps us to delineate between advertising that draws broadly on behavioural research and nudging. So far, I have suggested that marketing often uses methods of influence that exert greater control over individuals than nudges however, some adverts exert almost no control, relying instead on rational persuasion (see fig 2. for taxonomy of influence). Advertisers carefully study the target audience of their campaign; unlike nudges that tend to rely on common heuristics and biases, they know that sometimes certain audiences cannot be easily steered or manipulated. For this reason, they may sometimes opt for straightforward persuasion, detailing the specific qualities of a product. This
type of advertising is actually very similar to boosting, in that it might develop individual’s capacity for understanding how certain products work.

Christos Panousis claims that “some website advertisements can be identified as nudges” (2016). As in the case of ‘rational persuasion adverts’ and boosts, there certainly is an overlap in chosen methods. Adverts often exploit common heuristics by using methods such as default settings, showing how many other people have purchased a particular item (conformity biases) and citing ‘well known’ events or ‘facts’ (availability heuristic). In such cases it is difficult to know whether influences can be labelled nudges or adverts. The easiest way is of differentiating is to ascertain the intention or purpose behind said influence.

<table>
<thead>
<tr>
<th>Type of influence</th>
<th>Degree of control</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choice elimination</td>
<td>Fully controlling</td>
<td>A precipitously removes the possibility to ( \psi ) from ( B )'s choice-set.</td>
</tr>
<tr>
<td>Compulsion</td>
<td>Fully controlling</td>
<td>A uses physical force to get ( \psi ) to ( \psi ).</td>
</tr>
<tr>
<td>Coercion</td>
<td>Fully controlling</td>
<td>A threatens to make ( B ) worse off if ( B ) refuses to ( \psi ).</td>
</tr>
<tr>
<td>Behavioural prod.</td>
<td>Substantially controlling</td>
<td>A makes it more likely that ( B ) will ( \psi ), primarily by triggering ( B )'s shallow cognitive processes, while ( A )'s influence preserves ( B )'s choice-set but is substantially controlling.</td>
</tr>
<tr>
<td>Disincentive</td>
<td>Substantially controlling or substantially noncontrolling</td>
<td>A increases the probability of getting ( \psi ) to ( \psi ), primarily by providing ( B ) with some monetary or nonmonetary benefit.</td>
</tr>
<tr>
<td>Incentive</td>
<td>Substantially controlling or substantially noncontrolling</td>
<td>A makes it more likely that ( B ) will ( \psi ), primarily by triggering ( B )'s shallow cognitive processes, while ( A )'s influence preserves ( B )'s choice-set and is substantially noncontrolling.</td>
</tr>
<tr>
<td>Nudge</td>
<td>Substantially noncontrolling</td>
<td>A induces ( B ) to believe willingly, or form the intention to ( \psi ), primarily by presenting her reasons to ( \psi ).</td>
</tr>
<tr>
<td>Rational persuasion</td>
<td>Fully noncontrolling</td>
<td>A induces ( B ) to believe willingly, or form the intention to ( \psi ), primarily by presenting her reasons to ( \psi ).</td>
</tr>
</tbody>
</table>

Figure 2. Taxonomy of Influence (Saghai, 2013a)

It should now be clear how psychological behavioural insights can be implemented in market advertising as well as in nudges and boosts. The use of these insights can differ in implementation with each technique, likewise the ethical implications vary. In the following chapter I aim to directly respond to my titular question: Under what conditions, if at all, can (psychological) strategic behavioural influences be justifiably used to shape people’s choices?

4. Justifying or rejecting (psychological) strategic behavioural influences

In this section I explore more deeply some of ethical issues associated with nudging, boosting and market advertising. Furthermore, I shall argue that the implementation of opaque or subconscious influences is not morally justified. This means that boosts, by nature of their transparency, can be ethically permitted and in some cases so might advertising and nudging. My argument contains three strands, each of which centres on the ethical need for transparency.
4.1 Sustainability

My first argument, that rejects covert or opaque behavioural influencing, is largely informed by Hertwig and Grüne-Yanoff’s boost dimensions: **reversibility** and **programmatic ambition**. These elements not only conceptually distinguish between boosts and nudges but do so in a way that makes boosting ethically acceptable, if not desired. My argument for sustainability refers to moral sustainability and individual autonomy. For individuals to develop moral capacities they must interact authentically with the world, forging ideas and beliefs that will inform other behaviour. The programmatic ambition of boosts is to develop these capabilities, by giving individuals the necessary tools. These tools are often ‘generalizable’ meaning they can be used in various situations. This also means boosts are not required so often. This is because individuals, who have engaged with boosts, should possess generalised developed capacities for rational decision making. Of course, not everyone is likely to accept offered boosts yet, due to their transparent nature, they provide an option that individuals may freely decide whether to take.

This notion of self-development and empowerment is not unique to boosts, pragmatic philosopher Dewey, believed that people should be permitted or encouraged to “take command of their own powers, so that they can find their own happiness in their own fashion” (cited by Pedwell 2017, p.84). This idea supports the idea of autonomy as individual agency, freedom of choice and self-constitution. It does so by allowing individuals the freedom to form their own conceptions of the good, additionally it does not hamper their abilities to strive in recognising them. This kind of freedom promotes its own ends whereas covert nudges and market advertising require continual maintenance.

According to the DAGMAR model, advertising intends to raise brand awareness and maintain it. In order to maintain awareness, brands often launch long running marketing campaigns. These campaigns are revived when brand awareness drops. This demonstrates the unsustainability of market advertising. Like many nudges it targets one specific behaviour and is therefore reversible.

Kalle Grill does not see reversibility as an issue; he cites an argument made by Sarah Conly who claims some paternalistic legislation may help to habituate moral behaviour (2014). To some extent this argument seems valid since legislation does appear to reinforce social norms i.e. speeding is illegal : speeding is dangerous and morally wrong. However, laws can arise as a result of social norms too, i.e. changing attitudes towards homosexuality : homosexuality became legal. There is clearly a relational interplay between laws and moral norms however, legislation acts as an open statement that can be subject to scrutiny and debate. The belief that laws should be open is
sometimes described as ‘deliberative public reasoning’, furthermore it is essential to civic republican conceptions of liberty (Lovett 2018).

When new laws are created in democratic countries the general expectation is that the reasoning behind such change is made transparent. If legislation does help to habituate certain moral behaviour it is likely because individual decision makers accept the reasoning behind it. Accepting or understanding reasoning is primarily a function of the reflective cognitive system. As discussed in the section on attitude, system 2 processing enhances attitude strength. The argument for habituation does not apply to nudges, largely because nudges exploit system 1 processing, therefore in situations where choice architecture is removed, individuals’ behaviour is likely to revert. Where influences are transparent, individuals receive greater opportunity to engage their reflective processing systems. This is good, not only because it reduces the likelihood of individuals becoming the instruments of another’s will but because it reduces the power placed in the hands of unelected choice architects.

My argument of sustainability deals primarily with the concern that non-transparent behaviour steering undermines individual’s abilities to develop their own capacities for rational choice. Where individuals lack this capacity, choice architects may feel compelled to steer decision makers again and again. Where individuals are unaware of how they are being influenced they are denied the opportunity to resist, furthermore, opportunities for cognitive development are removed. The removal of these opportunities creates individuals who might be in greater need of steering (since they lack the capacity to behave rationally in many situations) thus choice architects perpetuate their own need.

4.2 Choice Architects

Since anyone can be a choice architect and choice architects can decide whose ends they wish to satisfy with behavioural influences (Thaler and Sunstein 2009), there is a great deal of scope for ethical issues of manipulation, corruption and control. In this section I shall use these issues in support of my claim that behavioural influences must be transparent.

4.2.1 Manipulation

Generally speaking, an influence is thought to be manipulative if it exploits the cognitive weaknesses or capacities of an individual (Vugts et al. 2018). This is clearly the case in some non-transparent nudges and many adverts. Nudges and adverts may employ manipulative method such as framing to steer individual behaviour towards chosen outcomes. The example of framing described in Nudge explains how framing in healthcare can significantly influence individual’s
decisions about certain treatments. Thaler and Sunstein explain that the phrasing of statistics utilizes behavioural insights, to predict and steer patient responses. If patients are told “ten in every hundred patients who had this treatment are now dead”, they will feel alarmed and far less likely to undergo treatment than if they were told “ninety out of one hundred patients have survived since having this treatment”. Many would argue that decisions as important as this should emphasise patient autonomy (Hammilton 2018). They could do this by presenting facts in as neutral a way as possible. I think the importance of a decision heightens the ethical need for transparency, however transparency is necessary in all manner of decisions. Choice architects are not privy to the details of each individuals’ lives, for this reason they are not in a position to weigh the importance of a particular decision. As Feinberg argues, fecund options cannot be removed if we wish freedom of choice to remain intact. Since we cannot know the long term or knock on effect of a behavioural influence they should be restricted to non-manipulative means.

4.2.2 Corruption
As the BIT director Helpern suggested, governmental nudges often require ‘tacit acceptance’. Kosters and Van der Heijen point out that this acceptance or public trust is not only expected of nudge-policies but any “government-led governance intervention” (2015, p281). This may well be the case, but it does necessarily provide and excuse for non-transparency. Perhaps nudge-policies serve to highlight that governments require more transparency. Civic republican philosophers often argue for “contestatory democracy” (Pettit; Maynor cited by Lovett 2018). This notion of democracy emphasises the needs of individual citizens to be offered opportunities to contest governmental decisions. Non-transparent policies clearly defy this notion of democracy and may therefore be seen as impermissible by republican standards. Governments that operate in covert ways seem even more liable to corruption.

4.2.3 Control
Civic republicans seem to reject non-transparent influences in a number of ways. Paul Hammilton writes A Republican Argument Against Nudging and Informed Consent (2018). It is important to understand how notions of liberty vary, thus when Thaler and Sunstein argue nudges are liberty preserving, we might first consider what is meant by liberty. Civic republicans have a rather stringent understanding of liberty as non-domination. Domination exists when political structures allow for individuals to be arbitrarily interfered with (Lovett 2018). Influencing another’s behaviour might count as arbitrary interference. Hammilton explains how doctors might dominate patients when they arbitrarily interfere with them by nudging their decision regarding future treatments. This could be done by framing information in certain ways.
Not only does our understanding of freedom inform how nudges may be ethically perceived but our understanding of many principles including autonomy. Hammilton refers to the arguments of Cohen who describes the conflict between beneficence and autonomy in biomedical ethics. Cohen describes these duties in terms of “providing optimal care” and “respecting patient choice” (cited by Hammilton 2018, p276). In general individuals expect ‘optimal care’ when they visit their doctor, however their epistemic understanding might not match that of their doctor. Hammilton argues that revising the standard of optimal care to that which the doctor believes is best, puts doctors in a position of arbitrary domination. If doctors view a particular treatment as optimal, they may nudge or influence patients into opting for it. Another interpretation of optimal care might include the direct offering of all relevant information i.e. informed consent. This model incorporates patient autonomy into the notion of optimal care and thus attempts to satisfy both duties to beneficence and autonomy.

4.3 Where’s the harm?

My final argument addresses the claims of Thaler and Sunstein, that choice architecture is inevitable. I show that not only is choice architecture avoidable, it can be dangerous. For this reason, unless it can demonstrably improve individual wellbeing it should be avoided.

Thaler and Sunstein reiterate throughout *Nudge* the inevitability of choice architecture. They argue that like all architecture there must be an element of design which serves a purpose (2009). They do however concede the point that some choice architecture is unintentional, whereas in the case of most nudges a clear behavioural outcome is intended. Grill notes that in fact choice architecture is not inevitable, since Thaler and Sunstein themselves write that, however pointless, choice design could be randomised. Since choice design could be randomised; designed with purposes other than to influence behaviour; or entirely without intention, it seems clear that the argument for inevitability falls short.

Since architecture is not entirely inevitable as a strategic tool for influencing behaviour, it is possible to argue for its ethical impermissibility however, I am sympathetic to the belief that it could be used to promote good outcomes. I think boosts or the ‘honing of new tendencies’ are possible solutions to the question of how choice architecture might bring about good (Pedwell 2017, p.88). The key element that separates these techniques from those such as nudges or advertising is the stringent belief that they must be transparent.

The need for transparency so far has relied on arguments that promote autonomy. Understandings of autonomy may however differ; for this reason, I now consider why covert influences could be
dangerous and therefore ethically impermissible in a more practical sense. Alluded to in the previous chapter was the concern that individuals might not develop moral character if continually guided in decision making. Luc Bovens describes individuals who are easily influence as ‘fragmented selves’ (2009). Bradbury et al. point to neoliberal ideology that works in harmony with nudging to create policy, as a potential cause for redefining ‘rationality’. They believe, “Rather than being liberated to act rationally, ‘rational’ subjects are constructed in a game of incentives and threats” (2012, p.250). If we are to accept the concerns of Bovens and Bradbury et al., the future for humanity’s morality does not look bright. Bradbury et al. believe that nudges provide the perfect way for neoliberal government to construct ‘rational’ subjects. Subjects are constructed through manipulative measure, deemed to be rational by those who design choice architecture.

Neoliberalism incorporated the private sector into governmental functioning, this muddies the waters in terms of accountability. Furthermore, many of the rehearsed techniques that can be used to influence individuals’ choices, come from market advertising (Whitehead et al. 2012). As previously discussed, advertising can be used to raise consumer awareness however, some advertisement reinforcement models have a slightly different emphasis. Jones distinguished between ‘weak’ reinforcement models, such as DAGMAR and ‘strong’ theory (cited by Heath 2007). He sees advertising not merely as reinforcement of brand awareness, but as a creative way to insinuate information; capable of drawing in new customers through persuasion. He believes it can persuade “apathetic and rather stupid consumers (by) the use of psychological techniques that destroy the consumer’s defences; in some cases, these techniques are not even perceptible to the conscious mind” (ibid, p.91). The belief that effected consumers are ‘stupid’ appears to serve as Jones’ justification for the use of subconscious or opaque methods. This attitude is reminiscent of the heuristics and biases research program that has been criticised for treating humans as ‘cognitive cripples’ (Edwards cited by Grüne-Yanoff and Hertwig 2017). Theories based on beliefs that humans are irrational or stupid seem often to use this as groundwork, upon which they justify autonomy violating principles or make claims that they are promoting individuals’ best interests.

Removing or steering individuals’ choices does nothing to enhance individuals’ intelligence or capacities for rational decision making. In fact, the purpose of nudge policies according to Bradbury et al. is to create subjects who adhere to conceptions of the good as defined by choice architects (2012). I believe manipulative methods of nudging and advertising are inherently dangerous not solely because they hamper individual’s autonomy and freedom but because they hamper the development of crucial traits.
5. Conclusion

In conclusion, I have argued that non-transparent behavioural influences that employ behavioural insights are not ethically justifiable. This is because they exploit human cognitive weaknesses without providing reasons that might address or correct said weaknesses. Individual decision makers are denied the opportunity to develop rational capacities when influences are non-transparent. Furthermore, non-transparent influences often violate individuals’ autonomy in deciding what choices would further their own best interests.

Organizations such as the BIT operate with limited accountability, effecting government and private-sector policies on an international scale. These organizations, if operating in a fully transparent manner, could be more easily monitored and therefore subjected to contestation. This would enhance individual liberty, as understood by civic republicans. Since global organizations seem to hold increasing power, there must be an appropriate response to ensure they are held accountable. Transparency is essential to this process and as such must be considered an ethical standard or condition.

Human beings must not be thought of as subject to be constructed, rather individuals with the capacity to decide on their own conceptions of the good. Boosts and perhaps transparent educative nudges can be implemented as tools that might enhance individual wellbeing (as decided by the individuals themselves). Market advertising is subject to legislation that limits its potential to manipulate consumers. Were individuals empowered to enhance their capacities for rational cognition, market advertising would perhaps provide less of a threat to those who had previously been seen as ‘stupid’ or irrational.
Bibliography

Books


Theses

Andersson D, 2016, *Deciding Fast and Slow: How Intuitive and Reflective Thinking Influence Decision Making*, Linköping Studies in Arts and Science Dissertation No. 698, Linköping University, Linköping


Articles


De Brigard, F (2010) ‘If You Like It, Does It Matter If It’s Real?’, *Philosophical Psychology*, vol. 23 pp. 43-57


Websites and online articles


