



**Œconomia**

History, Methodology, Philosophy

6-4 | 2016

Economics in the Shadow of the First World War

---

## Nudge: How Far Have We Come?

Magda Osman

---



**Electronic version**

URL: <http://oeconomia.revues.org/2490>

ISSN: 2269-8450

**Publisher**

Association Œconomia

**Printed version**

Date of publication: 1 décembre 2016

Number of pages: 557-570

ISSN: 2113-5207

**Electronic reference**

---

The text is a facsimile of the print edition.

## Nudge: How Far Have We Come?

Review essay on:

Cass Sunstein, *Simpler: The Future of Government*, New York: Simon & Schuster, 2013, 277 pages, ISBN 978-147672659-5

Cass Sunstein, *Why Nudge? The Politics of Libertarian Paternalism*, New Haven: Yale University Press, 2014, 208 pages, ISBN 978-030021269-3

Cass Sunstein, *Choosing Not to Choose: Understanding the Value of Choice*, Oxford: Oxford University Press, 2015, 240 pages, ISBN 978-019023169-9

Cass Sunstein, *The Ethics of Influence: Government in the Age of Behavioral Science*, Cambridge: Cambridge University Press, 2016, 234 pages, ISBN 978-110714070201-6

Richard Thaler and Cass Sunstein, *Nudge: Improving Decisions about Health, Wealth, and Happiness*, New Haven: Yale University Press, 2008, 320 pages, ISBN 978-014104001-1

Magda Osman\*

### 1. Nudge: The Pitch

Suppose that a government's social policy agenda is to promote a healthier, wealthier, and happier population. The problem is that our psychological makeup means that we find it hard to behave in ways that consistently meet these objectives. Over half a century, a steady accumulation of social scientific research shows that people are susceptible to cognitive biases, that they fail to exercise self-control when it really matters, give in to their emotions too easily, and make snap decisions without thinking hard about the consequences. Together, the frailties of being human mean that the population often makes poor personal, social, financial, environmental, occupational, and

---

\*Queen Mary University of London, m.osman@qmul.ac.uk

health related choices. Ultimately, all of this contributes to poor mental and physical health, which in turn leads to unhappiness. So, how does the government intervene in such a way as to improve the decisions the population make around health, wealth, and happiness, while at the same time preserving freedom of choice?

The typical tools of the policy maker's trade are mandates, bans, taxes, fines and financial incentives. In other words, the policy maker's tool kit involves inducing the population to make "good" choices by offering a carrot (i.e. reward), threatening a stick (i.e. punishment), or a combination of both.

Since publication of Thaler and Sunstein's (2008) groundbreaking book *Nudge: Improving decisions about health, wealth, and happiness*, policy makers can now add another tool to their kit: nudges. "Nudge is any aspect of the choice architecture that alters people's behaviour in a predictable way *without forbidding any options or significantly changing their economic incentives*" as defined by Thaler and Sunstein (2008, 6).

What does this mean in practice? Nudges, in the broadest of terms, are behavioural change techniques. These are methods that are designed to guide people to make better choices (as they themselves would judge as better),<sup>1</sup> by exploiting insights from psychological research. That is, if the policy maker knows what motivates people, how they process information, and how they make decisions, then it is also possible to construct the options that people face in a given context in such a way as to make the "better" option easier to choose. Hence the term behavioural change technique, because it uses psychological research in order to create changes in behaviour at a population level.

Two terms that Thaler and Sunstein (2008) introduce as part of the language of nudge are *choice architecture* and *choice architect*. Choice architecture means the way in which the options are presented to people in a given context when they make their decisions. We'll examine an example of this shortly. The choice architect is the individual/organization that fashions the options in a given choice context, in order to promote one option being more likely selected over another. In addition, Sunstein has developed a conceptual and ethical frame-

---

<sup>1</sup> The issue of what is "better" is loaded for two reasons. First, there have been several discussions regarding whether it is possible to evaluate nudged options as rational, or even optimal (e.g., Gigerenzer, 2015). Which has since led Sunstein (2016) to propose that a useful criterion by which the nudged choice ought to be evaluated as the "better" choice, should be on the grounds of what the individual judges as better. Second, this same argument is also used as a way to justify nudges on ethical grounds, because, however the nudge is implemented, the individual that has been nudged can judge whether their behaviour has been improved as a result of the nudge, if it has, then the nudge is justified. Put simply, individuals determine the desired "ends", and the ends justify the means.

work that helps justify the nudge programme, which he describes as *liberal paternalism*. Nudges aim to increase the likelihood of the “best” option being selected, but crucially they do this without eliminating freedom of choice. In other words, an authority might decide what choice behaviour is best, but rather than force that behaviour onto people, instead, people are guided towards making the “better” choice themselves, and can always still choose a “worse” option. This approach solves several major problems that a policy maker faces, the most obvious of which is how to “encourage” people to make good lifestyle choices that benefit them and society, while at the same time preserving their right to choose what to do (within the legal framework of the law of that country).

In order to gain a better sense of what these terms mean, and how they apply to real world situations, let’s now consider a typical example of a nudge: Automatic defaults. There is a long history of automatic defaults being used by public and private institutions in many different contexts (e.g., organ donation, pension schemes, health insurance, phone tariffs, energy providers) (Sunstein, 2015). Usually what happens is that the institution, let’s say in this case it is the government, pre-select an option that is deemed the “better” option, and it enrolls the population onto an organ register system. So, to clarify, the choice architect is the government, and the choice architecture is the construction of the options (i.e. donate, not donate), and the nudge is the pre-selected option “donate”. The liberal paternalistic justification here is that people should remain on the register if they have a preference for doing so, but that if not, then they can actively choose to “opt-out” of the organ donation register.

Why might this nudge technique work? Thaler and Sunstein (2008) claim that in this context there are emotional barriers to thinking about, and then actually doing something about enrolling on an organ donation register. So, by pre-selecting the option to donate one’s organs, the government increases the number of people on the organ donation register (which in most countries is still lower than the demand for organ donations), and takes away the mental effort and emotional pain. At the same time the nudge technique also exploits people’s inertia or “status quo bias” (Thaler and Sunstein, 2008), so that they are more likely to be locked into the defaulted option because there is a small cost in time and effort to “opt-out” of the register.

The essential message that is promoted by Nudge advocates is that nudges, just like the example given here, are simple to implement, cheap (relative to typical regulatory policy tools), and, are ultimately choice preserving because people are never prevented from choosing the non-nudged choice, only that it is harder for them to do so. More to the point, nudges can help promote a healthier, wealthier, and happier population.

## 2. Nudge: The Issues at Stake

As is the case with any successful idea, there are many commentators that will likely take issue with it. Nudge is no exception. In fact, given the popularity of the nudge programme, and the fact that it has been incorporated into many governments across the world, the work of Thaler and Sunstein has attracted close scrutiny. The occasion for this review is Sunstein's (2016) most recent book on nudge *The Ethics of Influence: Government in the Age of Behavioral Science*; a detailed discussion of the ethics behind nudge. The aim of this essay is to use Sunstein's (2016) recent book as a start and end point to a discussion of the evolution of Sunstein's views on nudge following his pioneering work with Thaler (Thaler and Sunstein, 2008).

Returning to the example discussed in the previous section, several questions on the theme of ethics have been raised. Is it ethical to automatically enrol people on an organ donation register, which promotes a decision that society benefits from, but that essentially relies on psychological trickery? On what criteria should we judge whether or not a nudge is ethical? When is a nudge designed by government simply regarded as having a benign influence on behaviour and when is it deemed coercion? These are the starting points for the first chapter of Sunstein's (2016) latest book in which a set of conditions are introduced that nudges fulfil on ethical grounds. The following four chapters then go into considerable depth on matters of welfare, dignity and autonomy, with a particular focus on Mill's Harm principle: "The only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others." (1859, 21) The reader is presented with many examples of nudges that have been implemented in a variety of ways and in a variety of contexts to help illustrate where and when state interventions are more or less ethical. To help support the arguments being laid out earlier in a later chapter Sunstein summarizes findings from several recent surveys designed to gauge public opinion on nudges. As a further defence of the thesis of the book, which is that nudges are indeed ethical, the penultimate chapter applies many of the criteria for evaluating the ethicality of nudges to show that they are ethical in the context of encouraging collective action to protect the environment. The final chapter rounds the book off by defending nudges by contrasting them to mandates, with a postscript titled "A very brief recapitulation" in which Sunstein's uplifting concluding point is that respecting freedom of choice has helped many nations work towards their ideals, and as they continue to do so, nudges are essential to fully realising their ideals.

What follows in the next section of this review essay is an examination of Sunstein's earlier proposals on the nudge programme. The aim of this section is to use the progression of his ideas up until his

most recent book as a way to understand how the nudge programme has been developed, and how the nudge thesis has changed. Three questions that constrain the focus of the review of Sunstein's earlier books are: 1) What are the distinctive characteristics that constitute a nudge? 2) If one is to go about constructing it with the view to implementing a nudge, what does a choice environment look like? 3) What makes nudges ethical?

### 3. Nudge: The Justification for the Nudge Programme

The point to remember about nudges, and what makes them so compelling, is that, assuming that they actually do work, the way in which they are supposed to work is by making the "better" option (e.g., a healthier food option, a pension plan, a fuel efficient home energy provider) easier to select. Crucially, this is done without having to significantly change the pay-off structure (i.e. without financially incentivising "better" options). All the hard work in generating a change in choice behaviour is done by utilising psychological processes and directing them to help people make choices that, in the long run, are good for them and society. These are the defining properties of nudge as presented in Thaler and Sunstein's (2008) definition of nudge.

So, perhaps anticipating an ethical critique, a key message established right at the start of the authors' first book on nudge (2008), and repeated throughout, is that nudges are unavoidable, they are everywhere in life, and that they are never neutral. Whenever any of us attempt to change the way we present information in such a way as to influence other people's choices, we are acting as choice architects—nudgers. Given how pervasive nudges are, there is little reason to take issue with them on ethical grounds. Any generalised ethical critique levelled at nudging problematises something all of us do on a regular basis.<sup>2</sup> There isn't much more development of this argument in Thaler and Sunstein's (2008) book.

---

<sup>2</sup> Well, it is true that setting everyday behaviour against some standard of absolute "purity" is rarely useful. In this example, setting Nudge against an ideal of absolute straightforwardness is clearly naïve (indeed, too easy a faith in straightforwardness can be dangerous: witness the controversy caused by the "straightforward presentation of the facts" in recent UK "deliberative polling" consultations, to take but one example). But surely we can acknowledge that a practice is pervasive without accepting that pervasiveness makes it morally neutral. Dishonesty and violence are also pervasive social phenomena, but they are constantly problematised, and controlled, by complex moral discourses; the ideals opposed to these problems, truth and peace, while never fully realised in the world, nonetheless give us conceptual reference points from which to launch discourse. The same must hold true for an ideal of *straightforwardness*—something like an ethic of integrity in nuanced, contested social environments where no agent can make simple claims to truth. Gladly, the need for the development of a more

What is the choice environment? While Thaler and Sunstein (2008) refer to the term choice environment often, they are never explicit about what the environment is. In other words, does it only refer to the options that people are faced with, in any given context, or does it also include the social, political, economic, cultural properties that also play a role in the way people make their choices? There is no answer to this question. In fact Thaler and Sunstein (2008) do not specify what the particular properties are of the environment in which nudges are implemented (i.e. where should they be used—and what are the best fits between the context and the nudges). They also don't say what the corresponding outcomes in the environment are that demonstrate the influence of nudges. In fact on the latter question, the rather glib response from Thaler and Sunstein (2008, 247) is that the evaluation of the success of nudges depends on their effects, which could be in terms of whether the nudges hurt or help people.

What are nudges? Also, what don't count as nudges? Towards the end of their book (chapter 14), Thaler and Sunstein (2008) list 12 examples of nudges which receive specific attention and discussion. In this chapter, the ranges of contexts in which nudges appear are hugely varied, they don't always occupy physical spaces, and the nudges themselves are mostly bans, or financial incentives. Almost all of the examples of nudges don't fit the authors' own definition of nudge, moreover, the choice environment is predominately economically structured, so that options are associated with outcomes that have financial rewards or costs.

For instance, two of the examples of nudge involve choice environments that are constructed in ways that involve financial inducement. One is focused on dissuading teenage mothers from having more babies, and the other is focused on dissuading drinkers from over-consumption of alcohol. In the former case the mothers get a dollar a day for every week they don't get pregnant which goes towards a college fund, and the drinkers are encouraged to use their own money to buy a drug that makes them vomit whenever they have a tendency to want to drink alcohol; the idea being that this negatively reinforces them to stop drinking. Both "nudges" are designed to induce behavioural change of abstinence, both involve the use of funds to help change behaviour, but what is the choice environment here? What is the specific nudge that is being implemented? Is it still the same in both examples? In fact, from the 12 nudges presented, it is very hard ascertain what makes a nudge a nudge, and what conditions are necessary in order to construct a nudge.

Suffice to say, if any researcher or policy maker wanted to consult THE guide book on nudges to decide what nudge to implement given

---

sophisticated *ethics of influence* is recognised, and at least partly met by Sunstein in the course of the next eight years since the seminal nudge book in 2008.

a particular problem, and how to construct the appropriate choice environment to promote “better” choice behaviour, the answers aren’t provided in Thaler and Sunstein’s (2008) seminal book. If anything the book provides a hazy, complicated and conflicting picture as to what the choice environment is, what a nudge actually is, and whether these nudges are effective in doing what they set out to do. What is clear is that a discussion of the ethics of nudges doesn’t feature significantly in the first book on Nudge, because that isn’t the primary objective of the book. The aim of the first book is to lay out the reasons for why nudges are a good alternative to conventional regulatory tools of government.

If we consult Sunstein’s (2013) book *Simpler: The Future of Governments* and then Sunstein (2014) *Why Nudge? The Politics of Libertarian Paternalism*, perhaps we will find more detailed answers to our three core questions? In both books (Sunstein, 2013, 2014), Sunstein makes the same strong claim as is made in the seminal book on nudge, which is that nudges can be found everywhere, and everyone at some point can be classified as a choice architect. This is beginning to sound like a pre-emptive strike against the looming threats against nudges on ethical grounds. If we wanted to know what nudges are, in both books (Sunstein, 2013, 2014) an explicit connection between nudges and financial incentives is made. In fact, Sunstein (2013) goes into some detail on a few key nudges, each of which seems to involve financially incentivizing people to select the “better” option (2013, 136). We see this again in Sunstein’s 2014 book, in which a host of examples of nudges are presented, but few of which actually adhere, in any strict sense, to the original definition of a nudge.

In Sunstein’s 2013 book one clue as to how to construct a choice architecture is mentioned: “The broadest point is that private and public institutions often fail because they make things difficult. Sometimes they make things difficult because they do not select simple, sensible default rules. Sometimes they create inadvertent problems because they list, and do not structure a wide range of options. Often the best approach is to make things automatic, so that if people do nothing at all, they’ll be just fine.” (2013, 128) Making options automatic is the job of defaults, and in fact the use of automatic defaults is deemed so important that Sunstein dedicates his entire 2015 book to the merits and pitfalls of automatic defaults.

Unlike the first book, in Sunstein’s 2013 book, we see a discussion of ethical matters concerning nudge, but really only from the point of view that the nudged “better” option is generally justified on ethical grounds because it serves the good of the individual, and the good of society. Crucially, what is deemed good is also one that the individual should also recognize as good, and this argument takes full force in his latest 2016 book. But still, the fact seems to be that Sunstein (2013) finds the matter of having to justify nudges on ethical grounds too

obvious to go into much discussion on it. For instance, Sunstein (2013) describes a case where a choice architect may construct an architecture that leads to an intended positive outcome (e.g. reducing asbestos in building materials), but also inadvertently generates negative consequences (e.g., increases the use of other cheap and toxic alternatives). This seems to be a concession that the motives of the choice architect matter when it comes to evaluating nudges, either on ethical grounds, or simply on the basis of achieved behavioural success. In fact this point is explored in his later books where he makes a distinction between a “means paternalist” and an “ends paternalist” (2014, 19; 2016, 54). This distinction is based on whether the choice architect is motivated to support an outcome by focusing on the conditions that will help to bring it about (means), or is motivated to prohibit an outcome by punishing it (ends).

Sunstein’s (2014) book starts much the same way as his earlier books, with a riff we are familiar with now. That is, choice architecture (whatever that is) is unavoidable, it is social, and that virtually everyone is a choice architect in some way (2014, 16). The same points are also used to make a general sweeping justification for the use of nudges by government.

Of all his books thus far, Sunstein (2014) is the most explicit in asserting the link between guided choices and the differences in the costs (i.e. personal harms which could be physical, psychological, financial) that they generate, and that the relative weight of these costs are to be assessed by reference to a graded notion of paternalism (2014, 51). This is where we see a dedicated discussion on ethics. He goes into great detail on Mill’s Harm principle as a talking point around different types of nudges, and their corresponding association with different shades of paternalism (2014, 52). That is, the harder the form of paternalism, the greater the costs to the individual in choosing the non-nudged option that is judged by the state/science as leading to harm (e.g., banning smoking in public places); his distinction between hard and soft paternalism is explored again in more detail in his latest 2016 book.

It might be fair to say that Sunstein (2014) is beginning to relent to the mounting scepticism from the academic community on the matter of nudges on ethical grounds. Not only does Sunstein (2014) concede that ethics matter, but also he concedes the potential limitations of nudges. He does this by suggesting that a one-size-fits-all strategy may not be the most effective approach to collective behavioural change (2014, 19), and that profiling an individual’s preferences and values might be a better way of effectively nudging people, because by doing this the choice architect develops a more nuanced understanding of what best serves the individual’s needs. He picks this point up again in his following book (Sunstein, 2015), but this time having had some time to think about things, Sunstein is critical of the

use of personalizing nudges (2015, 157-173), and spends a whole chapter discussing why that might be a problematic approach (2015, 257-173).

Sunstein (2014) dedicates a chapter to “the Paternalist’s toolbox”, which offers false hope for any choice architect that is embarking on constructing a choice architecture and seeks guidance on what it is and how to go about constructing it. Though perhaps, given that choice architectures are everywhere, this might be the reason why they are so hard to define. However, this section does provide the most explicit details thus far, on the nature of regulation—i.e. what types of nudges governments are likely to use, and the impact they are designed to have; these could be on the basis of specific outcomes on choice behaviour, changing preferences, and even influencing beliefs. To clarify this, Sunstein also outlines four classes of nudges: 1) those that affect outcomes without influencing beliefs and actions (e.g., automatic enrolment), 2) those that influence actions without influencing beliefs (a civil fine), 3) those that influence beliefs as a route to influencing actions (Educational campaigns), and 4) those that influence preferences as a way to influence actions, bypassing beliefs (graphic images on cigarette packets). While it looks like we are now getting closer to comprehensive answers to our three major questions, rather than go into more depth on the ideas developed in 2014, in the next book Sunstein (2015) abandons this, and instead presents a focused explanation on defaults.

Instead of giving people active choice, Sunstein (2015) presents four conditions under which a policy maker would be advised to implement automatic defaults: 1) when the context is confusing, 2) when people prefer not to choose, 3) when learning is not important, 4) when the population is homogenous along a relevant dimension (Sunstein, 2015, 18). Suffice to say, on careful inspection, some of the automatic defaults most commonly referenced by the nudge programme don’t meet these four conditions (e.g., automatic enrolment of pension plans, and retirement funds). As a reader we have come to be well versed in the general justification for nudges, in this case defaults. However, Sunstein (2015) does expand on the general sweeping arguments by claiming that most of the time people would rather avoid making choices, and don’t want to have to make a choice for every trivial matter (2015, 146). It is the responsibility of institutions to make informed decisions on behalf of individuals (e.g., computer settings, energy providers, health insurance, internet search engine settings, pension plans, phone tariffs). More to the point, people want to be freed up so as to have time to make choices about more consequential matters (2015, 207); presumably deciding on whether to donate one’s organs or not, is not a consequential matter. In addition, he argues that in the case of nudges, an automatic default does two things, it allows people the freedom to choose when they want to

choose, and not choose if they don't want to choose. This of course would seem perfectly ethical if people were actually making active choices one way or another (a = choose, b = abstain). But the whole point of defaults is that people do not choose and need not even be aware that there is (or was) a choice to make. Defaults by definition aren't neutral or symmetric with respect to the alternative options, and the preferences of the individual are presumed, along with their consent. Therefore, we are expected to place a lot of trust in the agents implementing automatic defaults on our behalf, which is a matter less serious when we have active choice. It is for this reason that Sunstein's example of the global positioning system (GPS) as a nudge (which he is using again in his 2016 book) (2015, 6) is misleading. The GPS pre-selects a route for us, but prior to even following the route, or not, we can choose to activate the GPS or not. When it comes to organ donation, and living in a nation that adopts an automatic opt-out system, people aren't actively choosing between choosing (donate/not donate) and abstaining. The final sections of Sunstein's 2015 book lay the foundations for further discussion in his 2016 book. The matter of coercion, and whether nudges are examples of coercive tactics (Sunstein argues that they aren't, 2015, 189-203), and the justification of defaults relative to mandates, are picked up again in his most recent book.

#### **4. Nudge: The Defense on Ethical Grounds**

In this final section we'll focus on Sunstein's (2016) latest book, and how it constructs the main thesis of the book, which is a defence of nudge on ethical grounds.

At the start of the book the question that Sunstein (2016) poses himself is "what are the ethical constraints on influence, when it comes from government?" (2016, 3). The short answer is: Welfare, Autonomy, Dignity and Self-government. At a later point he considers which one of these ought to be the ethical grounds on which nudges should be evaluated, and suggests that one should really only worry about welfare vs. autonomy of choice (2016, 62). Why might he say this? The clue to answering this comes from arguments he has made before in his other books, such as the fact that people do not always want to face making active choices (Sunstein, 2015; Thaler and Sunstein, 2008), and people shouldn't have to make choices over so many matters in which they could defer to experts (Sunstein, 2015); assuming of course that people can trust them (for a recent analysis of public opinion on experts see, Weingart and Guenther, 2016)

Along with this question, Sunstein poses several others about the ethics of nudge, and how nudges differ from coercion. His general response to these questions are that nudges are everywhere, and every environment in which we are primed, guided, directed, socially

manipulated, is a choice environment, and so we don't really have much of a legitimate basis on which to significantly worry about nudges since they are part and parcel of many choice situations we face. The problem is that this openly admits the fact that anything can reasonably count as a nudge (one also sees this on page 26 which includes a further list of diverse examples of nudges). From a scientific point of view, this presents the community with a deep problem, because it renders the nudge programme unfalsifiable. If one cannot really know what a nudge is, since everything counts as a nudge, how does one test the efficacy of nudges against other relevant non-nudge alternatives? Also, how can the scientific/philosophy community assess the ethics of nudge relative to other non-nudge alternatives, if it isn't clear what nudges are and what they aren't? A troubling consequence of this is when Sunstein tells us "Hitler nudged; so did Stalin" (2015, 25). These two individuals cannot possibly be examples of choice architects of nudges, assuming that nudges are free-choice preserving, by Thaler and Sunstein's (2008) definition.

Given that the book is a defence of nudges on ethical grounds, Sunstein presents 8 principles that defend nudges (2016, 18). The first principle is the one that states that nudges are everywhere and inevitable. The rest seem to be a collection of personal musings (principle 4: defaults can potentially promote autonomy, principle 5: dignity ought to be preserved if the choice architecture is good), confusing (principle 3: based on welfare it is a requirement that governments employ nudges on ethical grounds, principle 6: self-government can determine what nudges might be acceptable, reasonable, or should be prohibited) and statements of the obvious (principle 7: nudges can be bad if the choice architect has a malicious agenda, principle 8: manipulation is everywhere, and can be bad). Later, Sunstein delves into two central arguments that build a defence of nudges; these are, firstly: any critique of nudges on ethical grounds needs to be nuanced, examining nudges on a case by case basis, looking at the context they are applied in, the particular nudge being used, and the motivation behind applying it. There cannot be a single sweeping ethical problem with nudges because the nudge programme is made up of many different types of nudges implemented in different ways in different contexts (this is principle 2). The second argument is that a decision as to what makes a nudge ethical or not requires understanding of whether people consider that they are better off for being nudged, as judged by themselves; we have seen this latter argument before (Sunstein, 2014).

The first argument is a good one but for the fact that Sunstein is employing double standards. Recall that the first of his 8 principles is that it is pointless to take ethical issue with nudges because they are everywhere and the government uses them anyway. One can't really make a sweeping defence, and then criticize a counterargument em-

ploying the same tactic. The second argument is essentially an empirical one, and this is where Sunstein dedicates an entire chapter (Chapter 6) to discussing which nudges people like. Here he summarizes the results of several surveys conducted, including his own, that have been run on samples from multiple nationalities. The resounding pattern of results is that people show a strong preference for nudges that are deemed to be transparent (e.g., educational campaigns that inform people of details that help them make better decisions), as compared to nudges that are opaque and operate in ways that are not obvious to anyone engaged in the choice environment (e.g., re-ordering the food in a canteen so that the healthier options are presented first). Despite this pattern, Sunstein is rather generous with the interpretation of it. In chapter 6 he suggests that there is still a strong majority that would like nudges even if they aren't transparent, on provision that they work (2016, 146). There are two main problems with this chapter. First, it exposes significant issues with the nudge programme, by highlighting that it is in fact atheoretical. There is no framework or theory that is being used to explain the pattern of results discussed in Chapter 6, because there actually is no theory or framework (for discussion see Osman, 2014). Second, the surveys don't actually do what is necessary given the point of the book. The surveys do not report data on whether people consider themselves as better off for being nudged (as compared with typical regulatory tools), as judged by themselves. They are asked which out of a series of hypothetical situations they do have a preference for, and which they think are acceptable/not acceptable; this isn't necessarily getting at essential issues about people's own judgment on the value of nudges to them, and their trust in the choice architect's motives for introducing a particular nudge.

What the book doesn't do, that it should have done, is use the criteria it introduces on which to evaluate nudges on ethical grounds, and then systematically, chapter by chapter, take an example of a well-known nudge and go through and evaluate it, and demonstrate on those grounds how ethical, or not it is. Chapter 7 comes closest to this by focusing on nudges used to promote behaviours that are protective of our environment (e.g., recycling, reducing energy consumption, water consumption) and evaluating them on several criteria. What we can take away from Chapter 7 is that using defaults in the context of environmental protection is an effective means of positively changing behaviour where other regulatory tools have failed. Also, Sunstein suggests that active choice can be important, and in cases where it is not present, then one has to weigh the costs and benefits of making a particular option the automatic default. By which he means weighing the benefits of an automatic default against the cost of not delivering the experience of active choice, in instances where people do want to choose, but in fact (would) choose poorly. But, this all

seems to sidestep the point of the discussion, which is to answer the question: how ethical are nudges?

What is Sunstein's (2016) answer to this question? Ultimately it seems that Sunstein is of the view that nudges are indeed ethical, that they are useful, that governments are justified in employing them, and that when asked, the general public tend to like them, even the ones that they aren't aware of that are influencing their behaviour. Importantly, relative to mandates, people are still free to choose to ignore nudges. Of course people are free to ignore mandates as well, but the penalties are higher for doing this. So, what we can conclude here is that the ethical justification for the nudge programme boils down to the relative ease of choosing to ignore an option that is strongly signalled by a choice architect. But is that really the best solution that a programme of behavioural change has to offer?

## 5. Drawing Some Conclusions

Sunstein's work has helped to popularize behavioural science, by which he essentially means psychology, and his latest book also pays testament to this. More to the point, he uses work from psychology to showcase to a much larger audience some of the complexities of the rights of individuals to make choices and the role that the state plays (or should play) in order to encourage individuals to make choices that benefit society in the long run. Additionally, while it is the case that Sunstein (2015, 2016) acknowledges that active choice may be important, which implies that agency is of value, he doesn't say enough about why this might be the case, what the functional value of active choice is, and for that matter how that is realized at a fundamental psychological level.

Sunstein neglects the fact that there are key insights from psychology that demonstrate that the profound positive effects on mental and physical health, as well as decision-making behaviours, involve the promotion and strengthening of agency and autonomy over choice. This leads to actual not hypothetical changes in behaviour (for review see Osman, 2014). Agency and control are fundamental because they are the mental processes that contribute to reducing uncertainty of the future. Our mental representational structures are built on this premise, as well as our internal reward mechanisms, which drive actions that seek to assert agency and control. We cannot escape the fact that we are goal-driven, and this is precisely the result of a crucial belief that our actions can bring about outcomes that we set out to achieve; whether it is as mundane as making ourselves another cup of tea, or as complex as changing our career. These essential properties of our cognitive functioning are the very reason that active free choice seems so profoundly important to us on principle, and why it is that we feel aggrieved when it is subverted or denied. A

successful programme of behavioural change would do well to build its theoretical framework on these foundations, and along the way it can avoid the ethical tangles that the nudge programme is wrapped in.

## References

- Gigerenzer, Gerd. 2015. On the Supposed Evidence for Libertarian Paternalism. *Review of Philosophy and Psychology*, 6(3): 361-383.
- Mill, John Stuart. 1859. *On Liberty*. Oxford: Oxford University Press.
- Osman, Magda. 2014. *Future-minded: The Psychology of Agency and Control*. London: Palgrave Macmillan.
- Weingart, Peter and Lars Guenther. 2016. Science Communication and the Issue of Trust. *Journal of Science Communication*, 15(05)C01:1-11.