To what extent have the policy recommendations of the Behavioural Insights Team been in accordance with nudge theory?

By

Judd Ryan

S1619136

j.d.ryan@student.utwente.nl

Submitted in partial fulfillment for the degree of Master of Science, program Public Administration, University of Twente

February 2017

Supervisors:

Martin Rosema

Veronica Junjan

Word Count:17913

Summary: *The Behavioural Insights Team (BIT) was explicitly founded on the basis of nudge theory and was closely linked to it from the beginning. However, nudge theory has not contributed much to either the BIT's policies or governance in England and Wales. One explanation is that nudge theory is a limited policy tool, as critics claimed, but this has to be tempered with awareness of administrative explanations for disapplication. The ways in which nudge theory was disapplied are the subject of this paper. Nudge theory is defined, followed by a description of work of the BIT, before a comparison of the two. The first finding is that the BIT seems to have exhausted its policy ideas, implying that nudge theory and behavioural science is limited in policy applications. The second is that the disapplication of nudge theory reflects, in part, nudge theory's lack of coherence and definition. Implied in this is the suggestion that the strict moral criteria of nudge theory leaves relatively few nudges that can be made, and that there is a trade-off between effectiveness.*

Table of Contents

1. <u>Introduction</u>

Halfway through 2009 Oliver Letwin, one of the Conservative party's intellectual leaders, distributed a list of books that included '*Nudge: Improving Decisions about Health, Wealth and Happiness*' by Richard Thaler and Cass Sunstein (TS) to Conservative MPs, intending the theory it contained to form a plank of a future Conservative programme. The result was the Behavioural Insights Team (BIT), created to develop policy recommendations from inside the Cabinet Office. These recommendations were intended to be derived from behavioural science in the manner that TS propose, informing non-coercive psychological behaviour change techniques that enable an increase in social welfare. A term in power and an 'inside' book make now a good time to assess how this theory was implemented in government policy. This natural experiment can reveal much about nudge theory and its possible applications.

The appeal of nudge theory is the claim by TS (2008) that it can bridge the political divide by providing significant welfare gains without restricting choice or costing much money. It does this by tapping into behavioural science, using it to structure incentives in a manner that people subject to nudges choose behaviours that benefit them. This claim generated considerable excitement, and has even been conceptualised as a new mode of governance (Mols, Haslam, Jetten & Steffens, 2015). Oliver Letwin, the parent of the BIT, has been tempered in what he has said publicly, describing it as an "experiment", but one that was worth having for the £520,000 it was predicted to cost (Curtis, 2011). The BIT's debt to *Nudge* was made official as Thaler gave

evidence of nudge's efficacy to the sceptical House of Lords (Science and Technology Select Committee, 2011).

Few people outside of the nascent BIT and Cabinet Office expected the team to be a success when it was established in 2010, expected as it was to implement a theory that the press and policy-makers were dismissive of. Most were sceptical, even mocking, of introducing nudging to policy, and the claim that large welfare gains were possible through small policy tweaks that change behaviour. Within the policy community there were claims that Schipol airport's now famous flies in urinals was "as good as it would get" (Thaler, 2015, p. XII), while newspaper headlines implied the BIT was comedic in its absurdity (Hickman, 2011). Academia and important bodies such as the British Medical Association (BMA) were overwhelmingly critical, contrasting nudging negatively with traditional regulation, and arguing nudging simply preserved commercial freedom (BMA, 2012). Another criticism is that while *Nudge* is an interesting read, nudge theory, when practised, does not add much to the policy-maker's toolbox. This view is expressed with biting brevity in a question David Halpern (2015), the chief executive of the BIT, says he is frequently asked "isn't what you really do just about better communications?" (p. 151).

An idea that promises to revolutionise government and is then adopted by a government in the hope that it can will always be interesting, all the more so because of the widely divergent reactions it has generated. However, despite the claim of parentage, the BIT's policies diverge from nudge theory, sometimes radically so. It is necessary to draw a distinction between behavioural science and nudge theory. Nudge theory is fundamentally a technical and moral framework for the application of behavioural science in policy. In contrast, the BIT's remit

was to mine behavioural science and run trials for ideas that policies could be based upon. The distinction between a behavioural approach and nudge theory is a fine but necessary one. Nudge theory has an ethical dimension, and demands the maintenance of choice, that the policy is in the interests of the subject, and that the action is transparent. In this paper it is taken to be axiomatic that the moral criteria of nudge theory are sufficient in the abstract, and necessary if behavioural science is to be used in policy. Nudging is a form of manipulation that is both benign and beneficent, but if it slips from that it becomes manipulation. However, the experience of the BIT suggests that a strict application leaves very few policies available, and so the BIT has its own looser and ill-defined criteria that it uses.

The desire for levers to change behaviour had been a policy trend for some time, and in part was what TS were responding to in writing *Nudge*. On Halpern's part, he had long been an advocate for the use of psychology in government. Under his supervision the BIT has evolved into a 'skunkworks' unit, using any tools at its disposal and what it describes as an evidence-based empirical approach to behaviour change. The move away from the basic techniques of nudge theory again imply that nudge theory has prompt limitations. This view is further supported by the BIT's move away from giving policy advice altogether.

The first question to resolve is to define nudge theory, because despite having a nearly definitive and recent text to refer to it has been misunderstood and mischaracterised. This is done in the next chapter, followed by a literature review which together explain the misunderstandings of nudge theory, contradictions within nudge theory, as well as the response of writers to the founding of the BIT. A natural following question is whether these contradictions represent flaws that

make nudge theory inapplicable, and whether these flaws can be seen in the policies of the BIT.

Answering this will begin to reveal how useful nudge theory and behavioural science can be to policy-makers. The second issue is how the BIT has used or not used the basic technique of nudge theory. The next question is whether the BIT has satisfied the ethical boundaries that defines nudging insisted upon by TS, other writers and public opinion. These questions are answered in chapters five, where the BIT's activities are discussed, and six, where the policy recommendations are contrasted with nudge theory. Other possible factors in disapplication is another sub-question to explore. The BIT is now supported by a political consensus (John, 2013), but many of the concerns of academia have been borne out and the evidence suggests that nudge theory is both limited and poses a risk of abuse.

2. <u>Nudge Theory</u>

Nudge theory was developed by TS throughout the 00s based on growing behavioural science literature, which formed the body of evidence that showed market actors making predictably consistent non-Bayesian economic decisions. The desire to connect behavioural economics to policy happened in symbiosis with concerns that policy-makers had not been effective in changing behaviour, and that attempts to change behaviour were unpopular and restricted choice. Nudge theory posits that cognitive biases prevent people from behaving as Bayesian actors, or 'econs' in TS's terminology. The decision-maker, or 'human' in TS's terminology, is not making decisions not in line with their interests, but also not in line with what their preferences are when

unhindered by a bias. It is therefore legitimate to alter the structure of choices to promote the Bayesian choice. It does this not by removing biases, but instead by structuring choices to accord with the human's biases.

This combination of securing freedom of choice and providing welfare gains by concocting behaviour change leads TS to claim than nudge theory represents a true "third way" that solves "some of the least tractable debates in contemporary democracies" (p. 352). One need only see the recent resistance to bans on fast food in New York City to understand how strident people can be in defending their freedom of choice, even when surely aware that consuming 72oz sodas is detrimental to their health and the environment (McGonigal, 2012). With a nudge, everybody will maintain the maximal number of choices, but have those choices structured in such a way as to 'nudge' them towards the most beneficial or least harmful options.

This potentially places a great deal of power in the policy-makers hands, which necessitates TS placing ethical restrictions on nudging. TS named their theory 'libertarian paternalism' to distinguish it from coercive government action. Despite the name of their theory, TS are not libertarians, but the moral provisions within it mean, if properly applied, it is compatible with some visions of libertarianism. Nudge preserves autonomy and freedom of choice, and there is even a lengthy detour in *Nudge* to argue for marriage privatisation, but nudge theory does not promise to de-legitimise the political process. In this context 'libertarian' is a modifier to paternalism, intended to legitimise government intervention to change behaviour in areas where there is currently no popular support for intervention.

2.1. The Mechanics of Nudging

It is useful to distinguish between two elements of a nudge. The first is the mechanic involved that changes one behaviour into another. TS identify cognitive biases and flaws in heuristics as the reason for the apparent 'bad' decisions to be rectified by a nudge. These cognitive bias are a result of the automatic and unreflective mode of thinking that is uncontrolled, effortless, associative, fast, unconscious and skilled (TS, 2008, p. 20). TS's automatic and reflective systems have agnates in Kahneman's systems 1 and 2, with the intended outcome of nudging being to increase the amount of decisions made that would be made by the reflective mode of thinking, without having to engage system 2. The biases described within 'Nudge' that can lead to uneconomic decision-making include:

Anchoring, meaning over-reliance on early information given (p. 23);

Representativeness, especially misconceptions of chance (p. 26);

Overconfidence and optimism (p. 31);

Loss aversion, meaning that people over-value what they already possess (p. 33);

Status Quo bias, a causal factor in inertia (p. 34);

Biases arising from framing and apophenia (p. 35).

Nudges are not legitimate unless they are based on bias, because if no bias is present then it cannot be assumed that the behaviour is something the subject would wish to correct. Without this criteria, psychological techniques become just another tool for one group of people to enforce their will onto another, only through more insidious and

less overtly violent means, and achieved arbitrarily without the need for legislation. Nudges seek to channel these biases in positive ways rather than challenge them through persuasion or ignore them. In this way nudge goes beyond behavioural economics in trying to model how people will behave rather than simply observing it (Selinger & Whyte, 2012), and make the literature into a workable theory for policy-makers. Additionally, TS (2008) state that nudges are more effective when subjects are made to make decisions that are either difficult, rare, or both, for which there will not be feedback, and when subjects find it difficult to visualise the eventual outcome of their decision.

Increasing information is an ambiguous type of nudge. It is described as "highly libertarian" by TS (p. 189), who implicitly concede that some nudges are more intrusive than others. In doing so they give reason to doubt at least the relevance of their claim that every choice has an architecture behind it, and so every decision is nudged in some way. Information does not involve system 1, in contrast to what TS say is the purpose of nudging is to enable people to "rely on their automatic system without getting into terrible trouble" so that "their lives should be easier, better and longer" (p. 22). Relying on system 2, as is the case with providing information as a behaviour change mechanism, belies the point of nudge as it removes the cognitive specialisation that makes nudging effective. There are therefore two stools to the mechanics of a nudge: first to identify a bias that causes a sub-optimal behaviour, and second to alter the choice architecture to enable a better behaviour that still uses the automatic system.

2.2 The Moral Basis of Nudging

The metaphor of a nudge is one of friendliness, benevolence and gentility, but TS identify a need to defend the subjects of nudges against

abuse as would-be nudgers are potentially selfish and manipulative, and protected as technocrats. Protection is needed in spite of the inevitability of nudges; the entire world is choice architecture in which nothing is neutral. This claim about the inevitability of nudging is made throughout *Nudge* (p. 237, for example), and seems intended to silence concerns about nudging. However, that is quite different to deliberately structuring choices and TS knew this argument was vulnerable, and moral restrictions nudgers should place upon themselves to legitimise their policies are provided. Indeed, if some nudges can be less libertarian and more intrusive than others, and protection is needed, as TS acknowledge, there needs to be greater sensitivity the more system 1 is relied upon and less when system 2 is being engaged.

The first protection is a requirement nudges are used only when the desired outcome of the nudge is in accordance with the econ of TS's description, who make decisions using their reflective system. The key is that nudges must make the individual subjects' lives better "as judged by their own preferences, not those of some bureaucrat" (p. 10). The effects of nudge are not measured against a subjects' first order preferences, or "wishes", as Grill (2014) claims, but rather against a cooler and reflective idea of what is good for the individual being nudged. TS refer to their 'New Year's Resolution Test' (p. 73) as a way determining which behaviours can be judged as a subjects' preferences, and it is somewhat successful because the test is instantly familiar. If it is probable that the nudge aims at a behaviour that may be the subject of a resolution (e.g. to quit smoking), and implausible that a resolution may be made in the opposite direction (e.g. to begin smoking) a nudge is appropriate.

Inner econ, and not 'deliberative self' as some use, is an appropriate term to describe the subject TS are attempting to nudge, itself derived

from the term ('econ') used in *Nudge* to describe the rational self-interested actor of the classical model. Based on the clarity of TS's definition of subject preferences, as well as nudge theory's claim to libertarianism, it is evident that it does not permit choice architecture that sacrifices the individuals' welfare for gains accrued to an abstract collective, and certainly not to any other private interest. Further, nudges that appeal to econs are more likely to pass more people's new year's resolution tests than those that aim to achieve social goals. If the individual does not benefit (as judged by the nudgee themself) from the nudge, then any gains accrued are sectional gains, made by either an electoral coalition, politicians or the interests that capture them, charities or business, and not truly collective benefits. Nudges still achieve collective goals, but through the summation of individual ones, in a similar manner to the way the invisible hand of the market guides society to better outcomes. With this criteria mandating that nudges must be constructed in accordance with the wishes of the nudgee, nudge theory could expand the degree of autonomy people have over their own lives by making it easier to pursue their own aims, with cognitive specialisation reducing the effort necessary to achieve their own desired outcomes. Considering that choice architecture is ubiquitous, and that much of it pushes people away from their wants, choice architecture that reverses that could be liberating.

The fragility of this criteria raises obvious concerns. The 'econ' definition leaves space for interpretation as the subject cannot be asked, and even if they could be, their choices and opinions might be taken as the result of a bias. Further, it has to be assumed that preferences exist in the first place. The largest obstacle for policy-makers may be creating nudges that satisfy the diverse inner econs of whole populations. It is plausible that people may wish to begin smoking, or avoid becoming wealthier

when the population includes millions of people. This inability to decipher preferences, and their possible lack of existence, makes nudging far harder for government than is recognised by those who wish to integrate psychological approaches across government. The amount of nudges that the government can do is severely restricted by this criteria.

The impossibility of satisfying this criteria leads to the second moral criteria, that nudges need to be liberty-preserving (TS, 2008, p. 5). Nudges are necessarily avoidable, as TS express the fear that malign, biased or otherwise incompetent nudgers, and especially government actors, can misuse nudge theory (p. 10). The choice architecture built by a nudge must not only maintain maximal choice, but make avoiding the choice suggested by the nudge relatively frictionless, while not significantly altering incentives to the extent that they are essentially unavoidable, as is the case with Pigovian taxes (p. 6).

The preservation of choice can be read as both a moral matter, and as a desire to maintain the classical economic model. This definition has proven insufficiently precise, even for TS who offer as an example of a nudge Toxic Release Inventories (TRIs). This policy would significantly distort incentives for companies to reduce toxic releases (Hausman & Welch, 2010). It is not clear whether TS want economic freedom just for consumers, and favour regulation for business, or if they are genuinely interested in choice and classical economics. In other examples they show a disregard for the choices of those not directly subject to a nudge, such as businesses and taxpayers. They also have no problem with incentivising some behaviours, for example with financial rewards, but prefer not to disincentivise or punish behaviours. This is a contradictory position for TS to take given their argument that choice architecture is ubiquitous. It is actually a framing device in itself, and shows the how

easy it is to use psychology in a deceptive way. Further, all choice architecture does affect incentives in some way, and there is no way of clarifying how much incentives can be altered before it is said to be too much.

TS are acutely aware that nudging may be used for malignant purposes, to the extent that Thaler signs every copy of *Nudge* with the plea "nudge for good" (Hansen & Jespersen, 2013). The third layer of protection for nudgees is to call attention to Rawls' publicity principle, banning governments from nudges that they would not be willing to defend before electorates. Sunstein later judged this to be inadequate and insisted on transparency instead, saying that with "official nudging...transparency should be built into the basic practice" (Sunstein, 2014, p. 584).

Therefore nudge theory, despite the lack of clarity, has a moral character that restricts the range of policies available and matches the agreeableness of the nudge metaphor. Nudgers are faced with estimating the preferences of large bodies of hypothesised nudgees, deciding how much incentives can be altered before liberty can be said to have been restricted, as well as maintaining transparency and the effectiveness of nudges. There are other qualities that nudges normally have but are inessential. Given that part of the aim is to reduce costs, or at least not increase them, then it can be said that nudges need to be inexpensive. Even the financial benefits of the nudge outweighing the costs is not enough, because large costs, especially if they come at the expense of the taxpayer, have already distorted incentives. Another criteria is that nudges should appeal to both the left and the right of the political spectrum. Following on from the previous criterion, the reduced role of the state should satisfy the primary political aims of conservatives, while welfare gains should achieve the same for leftists. A further

characteristic of nudge theory is that it is innovative, promising, through the introduction of behavioural science to policy, new solutions to complex social problems.

3. <u>Literature Review</u>

The literature diverges on a few key questions that are important to understanding how policy-makers, and the BIT in particular, use nudge theory. The first is the definition, coherency and applicability of nudge theory, which some writers have tried to expand upon to make it more workable. The second question to occupy academics is whether or not governments, and especially the BIT, should nudge, focussing on both the ethics and efficacy of nudging. This has mainly centred on the effect on freedom, with academia in the UK mainly contending that it undermines collective freedom to deliberate and learn. This is important because nudging is undermined if it does not benefit the people it is intended to, or is not supported by them. The final issue is whether or not what the BIT has done so far constitutes nudging.

<u>3.1 Different Interpretations of Nudge</u>

The term 'nudge' is itself disputed. In this paper 'nudge theory' is preferred to TS's own term 'libertarian paternalism' (the nomenclature TS give to their theory, they use 'nudge' to describe the action) because that is the term that has been converged upon. The idea that TS's theory is either libertarian or paternalist has been widely rejected by writers. The conflict is between those who take a view of liberty as an absence of coercion (as TS do in 'Nudge', and as libertarians often do), and those who see liberty as a matter of individual autonomy (Hausman & Welch,

2010; Grune-Yanoff, 2012; Yeung, 2012). Likewise, advocates of paternalism reject TS's view that providing information is a paternalistic act as incorrect, and who insist that coercion is recognised as the defining feature of paternalism (Hausman & Welch, 2010).

Gigerenzer (2012) argues that libertarian paternalism and nudge theory are mistakenly conflated, and that nudging essentially means any non-coercive way of guiding behaviour. In this scenario, libertarian paternalism is one way of enacting nudge theory, which Gigerenzer seems to want to include all behavioural approaches. This is to be rejected because it provides policy-makers a canvas rather than a commission, and is probably the result of wanting to tease the meaning of nudge theory from its necessarily libertarian roots.

Some writers, such as Anderson (2010), have argued that TS defined the concept of a nudge too broadly, and that the examples they provide do not concur with nudge theory as TS themselves explain it. An important issue is whether the provision of information and attempts at rational persuasion are nudges. TS do include the availability, or unavailability, of information and how it is presented as a bias, but providing information does not use the subjects' bias in a positive way. By its prosaic nature, information does not necessarily alter a subjects' choice architecture at all. Hausman and Welch (2010), for example, argue that providing information is not nudging for this reason.

In contrast, Hansen and Jespersen (2013), using Kahneman's terminology, make a distinction between type 1 nudges, that are purely automatic, and type 2 nudges, that engage the reflective system. While this is more of an expansion of rather than an interpretation of nudge

theory, all nudges could be placed on a spectrum between the two, even if those closer to the type 2 end resemble either reminders, persuasion or bribery. TS wrote as if all nudges would be through the automatic system, but that they prescribed space to avoid nudges, as well as for information-based nudges shows that they expect some reflection. There is no reason to say that policies that engage the reflective system and ignore the automatic system are not nudges, as long as they overcome a bias by altering choice architecture, but it does belie the point, because nudges are supposed to work with cognitive biases. A theme in the literature is a desire to clarify and expand upon Thaler and Sunstein. Baldwin (2014) is another writer who argues that use of nudge theory has so far lacked theoretical clarity. He proposes three distinct types of nudges: those that enhance reflective decision-making; those that seek to create a bias towards a particular decision; and those that manipulate subjects, particularly by framing issue in a particular way.

The moral basis of nudging has been even less understood. The econ that TS wrote of became a 'deliberative self' to some (Oliver, 2013), a more permissive definition that provides another moral dimension to nudging. The result of this definition is to give nudgers a great deal more power, to decide what nudgees morally want and direct them towards that. What TS intend is a morally agnostic state in a libertarian fashion structuring choices so that people choose behaviours that they really want to choose. Others lament underlying selfishness (Barton & Grune-Yanoff, 2015; Grille, 2014), criticising nudge theory as promoting the well-being of the individual at the expense of collective goals. There is some ambiguity in *Nudge*, because examples TS provide, such as nudges designed to improve environmental behaviours, clearly aim at collective benefits. This clash between what TS say and the the examples they provide is the main contradiction in nudge theory. There

are collective goals discussed within *Nudge*, especially in the chapter 'Saving the Planet', but there are no instances in which it is suggested that collective interests supersede individual ones.

Hansen and Jespersen (2013) find the ethical clauses of nudge theory wanting, especially relating to transparency, and seek to create a clearer schema. Since then, however, Sunstein published a seperate paper to call for transparency. There is no answer to the question of how a nudge can be both effective and transparent, a widespread criticism (Hansen & Jespersen, 2013).

3.2 Should policy-makers Nudge?

Academia in the UK and elsewhere varies between centre left support for nudge, such as from John (2013) and De Ridder (2014), who offer unreserved endorsements of nudge, and the critical left, who reacted to the BIT by creating a narrative of Conservative malfeasance. This discourse represents the BIT as a cheap and ineffective token gesture in the midst of austerity policies, in place of a solution to the social needs of the UK. Pre-eminent is the idea that social change is best sought through collective deliberation (Goodwin, 2012; Leggett, 2014; Martin, 2015), while nudging is claimed to reinforce neoliberal norms such as individual responsibility and consumerism (Brown, 2012; Corbett & Walker, 2013; Leggett, 2014).

Behavioural economics and nudge theory are in this view a necessary reinforcement for a neoliberal order in crisis. Bradbury, McGimpsey and Santori (2012) pessimistically conclude that nudge theory just shows neoliberalism's malleability, maintaining the status quo while shifting its basis from rational man to irrational man. Corbett and Walker (2013) see

the BIT and nudge theory as an attempt to reinforce neoliberalism by making people increasingly "rational actors in a market scenario" (p. 459). This view imagines neoliberalism as the Learnaean Hydra, simultaneously frightened by its unappealing qualities, while in awe of its resilience.

Despite this, these same writers seem hopeful that the classical model has been undermined by behavioural economics. This is mistaken, nudge theory is built on the idea that there is a Bayesian self, an inner econ, beneath sub-optimal decision-making. Not even Adam Smith viewed economic agents as fully rational Bayesian maximisers of subjective utility, often meditating on what would today call bounded rationality (Coase, 1976). Nudge theory was not a response to anything other than a more nuanced understanding of the workings of the human mind. It was developed in the early 2000s, while most people felt no need to defend capitalism, by Thaler and Sunstein, one of them a Chicagoan economist, and adopted most notably in public administration by the British Conservative Party. There is no reason to think that either believes capitalism to be in crisis or based on flawed assumptions about human nature.

3.2.1 Ethical Concerns

As discussed before, there is concern about the ability, or intention, of institutions to implement nudge theory transparently, to secure freedom from manipulation and about the implicitly technocratic nature of nudge theory. Mostly, however, writers reiterate concerns that TS answered in *Nudge*. Distinct to the UK is an unfalsifiable line of criticism focusing on the BIT supposed usurpation of moral capacities, risking stunting subjects' moral capacity and causing infantilisation. Deliberative democrats (such as Goodwin, 2012; Leggett, 2014; Martin, 2015) see

the BIT as a way of bypassing existing deliberative procedures that have largely failed in Western democracies. This then reduces possibilities for active citizenship, stunting opportunities for personal and collective development, claimed by Martin (2015), who writes from the Habermasian perspective in which moral norms and participles of justice can only be agreed upon after active and open discourse. Without arguing for the disbanding of the BIT, Martin (2015) believes the deliberative deficit can be overcome if there are 'moral educators' who take an active role in public moral justification, for which the end point appears to be social democracy. Without this role, students who are nudged towards healthier diets and households nudged towards lower carbon emissions do not engage with the ethical dimension of their decisions. This amounts to a claim for the superiority of 'think' approaches. Other writers argue that nudge theory, when implemented, may well provoke greater reflection than is otherwise the case (Baldwin, 2012), which both undermines expectations about nudge and the deliberationist argument.

More than that, the deliberationists are concerned that citizens do not engage with their responsibilities to the collective or contemplate the moral possibilities of collective processes and decision when nudged. The effect is a 'de-moralising' of the public sphere as the privileging of the economic rationality of TS reduces all decisions to individual ones. These claims are echoed by Leggett (2014), who sees the BIT as the culmination of neoliberal governance in the UK. Against this background Leggett (2014) makes a case for a social-democratic behaviour change state. In doing so he argues for the importance of behaviour change on an individual level to solve the 'wicked problems' of the day, and that it will inevitably have a policy dimension. The BIT is seen as an ineffective way to tackle these problems and an attempt to solve social problems in

a cheap way that suits the Conservative's austerity agenda, and to solve what he sees as neoliberalism's "legitimacy crisis" (p. 8), echoed by (Goodwin, 2012), and emphasising self-responsibility at the expense of collective action (Martin, 2015).

Leggett (2014), therefore recognises a practical value to nudging but prefers "participative versions of the democratic ideal" (p. 10). These ideals are supposed to allow for development, while nudging results in a fragmented self unable to participate in collective deliberation and sees citizens' preferences subverted for those of experts. The solution is then to politicise behaviour change, though the effect this might have on efficacy, given that nudging arguably relies on cloaking a nudge, is not discussed.

An understanding that the BIT approach is reinforcing norms of collective responsibility is lacking, and this reinforcement is happening despite nudge theory being founded on science that strongly supports the superiority of individual solutions. If the BIT attempts to intervene, as it does, in people's diets, consumption, health choices, their decision to vote or not, then it is not relegating responsibility to the individual. Accepting that people have some control over these behaviours is not to blame those people, it rather frees them of their presumed hypoagency, and the BIT actually sustains a norm of government intervention in the process. Given that Leggett (2014) and Martin (2015) prescribe social democracy as the outcome of their deliberative democracy, perhaps the reason the BIT's mode of channelling collective resources finds their disfavour is that it does not base the need for intervention on overcoming some apparent injustice. That people of lower socio-economic class are more likely to be obese is not presumed by the BIT or nudgers to be the result of capitalist despotism, or defects unique to that class, but instead

the result of how the neurological constitution of humanity interacts with the choice architecture it is presented with. Further, the result of a deliberation might well overthrow the norm of intervention, given the growing taste for advanced liberty, and that many individual interventions are unpopular.

Nudge theory is a nuanced, balanced and compassionate understanding of individual responsibility, based upon firm empirical evidence. Gigernezer (2015) attests that people can be Bayesian decision makers, and that behaviour that meets TS's New Year's Resolution Test correlates positively with education. This, however, does not disqualify nudging from being useful to both those who may change their behaviour through education and those who may not. Even very educated people may continue in behaviours that are harmful to themselves and may benefit from a nudge. Problem gamblers are very aware that gambling is harmful even as they are in the process of gambling. In this instance education has failed when a change to the choice architecture, such as the ability to self-exclude from gambling possibilities, may be successful. There is also no reason to believe the two are mutually exclusive, either.

3.2.2 Efficacy

A line of criticism is to simply mock nudge theory, and dismiss nudging's claim to uniqueness and the strength of its scientific foundations. Yeung (2012) argues that there is nothing new about nudging as a technique, just its pretence to a scientific basis, dismissing it as a "fudge". However, nudge theory is deduced from a well established behavioural science literature - which Yeung does not contest - so there is already much that implies its effectiveness. Most published trials come from government bodies, foremost the BIT, and have overwhelmingly attested to the

efficacy of nudge, though mostly on relatively small scales. There are some exceptions to this trend, for example Moreira, Oskrochi and Foxcroft (2012) who show the failure of personalised normative feedback in decreasing alcohol abuse among British students. Additionally, nudging has also proved less effective when applied to people of lower socio-economic classes (Kosters & Van der Heijden, 2015). These failures provide support for the BIT's approach, employing trials before every formal policy recommendations to provide the best chance of success in particular instances. More importantly, however, it casts doubt on the ability of either nudging or the BIT to reach particular populations, perhaps those most in need of assistance, or to solve complex problems.

Another concern is that positive effects may be ephemeral and that trials have only been able to demonstrate nudge's short term impact. The question of the effectiveness of BIT therefore becomes mainly one of whether the gains it purports to have made have endured and can endure further. This claim is made by writers such as Martin (2015), for whom the gains made are characterised as ephemeral due to their lack of effect on the norms of citizens. Similarly, Mols et al. (2015) argue that nudges are ineffective, based on the idea that social identity is more important to shaping behaviour in an enduring way.

Halpern (2015) indirectly answers these questions when he makes the case for the automatic system. What is important is "habituation" (p. 322), as people establish new "behavioural equilibriums" (p. 320) and new neural pathways are established. Furthermore, habits are mutually reinforced by social norms and by commercial interests adapting to new behaviours. This incremental, market-based approach is hoped to be

more effective where education, and basically all government-led approaches, have failed.

A theme in the literature is the contrast with regulation. If policies that restrict autonomy are more successful, then the case for nudging is undermined (Di Nucci, 2013). Grill (2014) endorses nudging, but alongside 'jolting', defined as structuring choice content as well as choice architecture. In this pale endorsement nudging is just another tool of government, not "immune to disuse, universal policy solutions, a third way in politics, or as having any such more extravagant properties that proponents have claimed for nudges" (p. 139-140). In contrast to nudging being undermined, the fear that nudging may supplant conventional regulation is increasingly being floated (Oliver, 2013). Indeed, it may be that nudging is more effective, less burdensome and more popular than regulation.

3.2.3 Public Opinion

Given that successful nudging depends on people's preferences, if people subject to nudges are averse to either the means or ends of a nudge, then nudge theory is in jeopardy. The idea of nudging can be confounded if nudges are not in accordance with the hypothesised inner econ, or if this inner econ does not exist either for an individual or across the disparate people united in being subject to the same nudge. Further, nudging will fall short of this criteria if nudgees reject nudges as affecting their autonomy of mind and thought, or if the nudges as means either contradict or outweigh their beneficial impact.

Selinger and Whyte's (2012) claim that people are uneasy being nudged and are conflicted about the nature of autonomy has mixed support.

Studies that have looked into attitudes of nudgees indicate instinctive support for nudges as an abstract concept, but that consumers become more sceptical when nudges become more intrusive, when the nudge may cost them money, and when the aim is to promote a collective benefit, using the deliberative self standard, instead of a personal benefit (Junghans, Cheung & de Ridder, 2015; Hagman, Andersson, Västfjäll & Tinghög, 2015). There is some evidence of worries of nudging's manipulative nature (though one response to this was to wish to remain ignorant of the nudge), and when children are the subject of a nudge (though some respondents believed children to be ideal subjects for nudges), and that nudgees with an individualistic world-view are more suspicious of nudging (Junghans et al., 2015; Hagman et al., 2015).

No respondents in either survey offered the explanation that choice architecture is inevitable, though some did concur when that argument was made to them, and several knew that similar techniques occur in commerce. There is some consternation on the legitimacy of deciding what options are better than others. Some critics point to the lack of evidence for the health benefits of slimness as well as increased stigmatisation of those unable or not willing to comply with behaviours judged to be good behaviours (Junghans et al., 2015). The limited literature on the reaction of people to being nudged therefore reveals that even small samples contain a variety of opinions and concerns, and confirms that nudging is a sensitive and complex policy tool.

3.3 Nudging and the BIT

A second debate concerns how prevalent the behavioural approach has become in Whitehall. Some argue that nudging has become the default option for public servants (Jones, Pykett & Whitehead, 2013), while others argue that heralding the 'psychological state' or 'nudgeocracy' is

too great a leap to make (Kosters & Van der Heijden, 2015). What this debate misses is that whether or not the BIT has become an important artery of policy of policy is secondary, what matters is that the content of policies have not changed greatly. Norms about the role of government have been reinforced, and nudge theory transgressed to the point that the BIT risks becoming "theoretically empty" and "gimmicky" (Oliver, 2013, pp 10, 14). Mols et al. (2015) make the point that much of what the BIT has suggested rests on persuasion, appealing to the reflective rather than automatic system, while Oliver (2013) notes that many of the BIT's policy proposals follow conventional economic logic.

Despite academia being opposed to the BIT and nudge theory, they have not provided any compelling reasons to disregard nudging. There are some reasons to suspect it may be ineffective, as with every policy, but none that match TS's arguments and evidence. Equally speculative are claims about the potential of deliberative democracy. Taking agency away from people would have the effect of stunting moral growth, but no more so than any other government policy, and that doesn't mean that nudges cannot be useful, or that deliberation will solve wicked problems. That criticisms came solely from the left is to the detriment of the debate, ignoring as it does popular concerns about individual liberty that is vibrantly debated in the blogosphere. The only research to address these concerns looked at the attitudes of those likely to be nudged, which revealed concerns that ought to restrict what policy-makers see as legitimate to nudge.

4. **Methodology**

This paper follows a simple structure: nudge theory has already been described, and in subsequent chapters the work of the BIT will be described and then that work will be compared to nudge theory. This was also the research design: first review the literature, then compile the necessary data from the publications of the BIT, and use that data to analyse how the BIT's work matches nudge theory. The most salient concepts of nudge theory that can be operationalised in a comparison with the work of the BIT was described in the second chapter, and summarised below:

1. Mechanistic Criteria

    i.   Responding to a pattern of harmful or sub-optimal decision-making (judged against the nudgee's inner econ) caused by a cognitive bias that impedes decision-making.

    ii. Seeks to use that cognitive bias itself in innovative ways that lead to better decision-making for the subject.

2. Moral Criteria

    i.   Must have an intended outcome that accords with the nudgee's inner econ.

    ii. Must not restrict choice or alter incentives to the point that they act as an effective barrier to choice.

    iii. Must accord with Rawls' publicity principle and be transparent.

After reviewing the background literature, the dataset that formed most of the research was the twenty-two policy publications of the BIT from its founding in 2010 until the 8th of July 2016, which contain forty-three distinct policy recommendations, both listed in the appendix. Initially it was fewer, but more reports were published as work progressed, and it was necessary to include all possible policy publications to get as large a sample as possible. Indeed, because even all the contents of these policy reports represents only a minority of the BIT's work, it was appropriate to search other sources as well, including media reports and other BIT publications. This was archival research and involved analysis of the content. Most the reports began with a clearly identifiable policy problem or a multiple of them and a broad intended outcome. Towards the chronological end of the dataset the reports have increasingly unspecific general recommendations and cease to refer to policy.

This is as clear and operationalisable explanation of nudge theory possible. However, the concepts remain unquantifiable, as taking a very strict definition would mean almost nothing meets nudge criteria. An example is that many nudges satisfy the inner econ of large portions of those subject to them, but not others. A value placed on that would inevitably be an arbitrary representation of an population preference. This restricted the study to being an exploratory one, but one that can still resoundingly establish the falsity of claims made about the nudge theory and its relationship to the BIT. That the reasons for this, especially the utility of nudge theory, cannot be firmly established is what means this paper can only be exploratory.

The data was supplemented by information from '*Inside the Nudge Unit*' by David Halpern and the BIT's academic publications. This provided background to the data, but it also revealed its limitations There is a bias

in this methodology as not all policy recommendations were published. Many recommendations were probably informal and are not available anywhere, so it is not possible to provide an exhaustive list of the BIT's recommendations. It is possible that the trials and recommendations with the least internal consistency and moral restraint would be more likely to be rejected, and so those recommendations may deviate from nudge theory the most. It may be the case that they were stifled by an administrator with an incentive to see it unpublished (John, Sanders & Wang, 2014).Reports, or parts of reports are published pre-trial, and so the potential to judge the success of and response of the public to the trials is hampered. One strength of taking a largely descriptive approach was that it allowed a large amount of data to be used, but the main limitation is that it cannot definitively answer why the BIT has recommended the policies it has.

5. <u>The BIT</u>

Founded against a backdrop of scepticism and mockery, the BIT since seems to have won over opposition parties, Whitehall, as well as a sceptical media. In doing so, it has lost the qualitative differences it might have had that could distinguish it from other tentacles of government, and has become divorced from its original purpose, even moving away from trying to make the welfare gains that is the ultimate end of nudge theory. The decision to appoint David Halpern, a New Labour apparatchik as the CEO should have served as a warning that conservative and libertarian concerns were not the basis of the BIT, but rather a focus on administrative efficiency was.

Halpern's book reveals a picture of a team looking for every opportunity to implement their ideas while being given short shrift. Presented as an inside book revealing the inner workings of a unique government body, it reads more like a triumphalist account of growing influence and power in the face of hidebound resistance, as well as a glowing account of its own success. Concluding with strident claims about behaviouralism's potential to transform governance around the world, *Inside the Nudge Unit* is best seen as an imploration for more opportunities to make policy, with a hint that the BIT has not been as influential as Halpern writes.

5. 1 Administrative Background

The BIT has roots in government prior to the publication of *Nudge*, going back to the behaviour change agenda of New Labour. In 2001 the Forward Strategy Unit was formed, later to morph into the Prime Minister's Strategy Unit (PMSU), which employed Halpern, and wrote 'think pieces' advocating for the use of psychological insights in policy (Halpern, 2015). The think piece 'Behaviour Change' published by the PMSU attracted press criticism for this advocacy of psychological techniques, and slowed the trend. The PMSU did not survive the transition from Blair to Brown, but Halpern (2015) credits the unit with pointing Lord Turner in the direction of choice architecture, culminating in the changing of defaults in workplace pensions from opt-in to opt-out.

The behaviour change agenda continued, however, just without input from psychology. An example is DEFRA's 'A framework for pro-environment behaviours' (DEFRA, 2008), which used leass controversial mechanisms such as education and persuasion. Other approaches that pre-date the BIT include a focus on citizen-consumer information, shown most notably in increasing dietary information presented on food labelling, and, more importantly, 'evidence-based

policy' (Parsons, 2002). Just as important were the internal dynamics of the Conservative Party, for whom there is no doubt that nudge theory was a decisive fulcrum. Halpern (2015) makes plain that dislike of big government was crucial, and that it was an attempt to attain welfare gains (or, as critics allege, mitigate losses) at little expense. Since 2010 the BIT has managed to survive past the two years it was publicly conjectured to exist, and has since grown from seven to sixty members, been semi-privatised and partnered with governments around the world.

The psychology-based behavioural approach gained a new impetus with a new government and with the meek and novel sounding '*Nudge*' providing the motif, the British press mostly mocked, rather than railed against, the BIT. However, the case can be made that the debt to nudge theory is a misnomer. The administrative background suggests another interpretation, that of continuation of the epistemological communities in Whitehall that used the Conservative desire for a smaller state to establish the embryo of a psychological state at the heart of government. There is a case that the policies the BIT pursued owe more to pragmatic Blairism than either Thaler or Sunstein or whatever stunted libertarian impulses Conservative leadership may contain. In this interpretation the overlap between nudge theory and the behavioural approach is not necessarily causal, but because of a common parentage in psychology. This can be rejected, because much of the language, arguments and imagery used by the BIT were taken from *Nudge*.

5.2 The BIT's Other Work

The BIT has progressively deviated from its original purpose of providing policy recommendations, indicating a desire to incorporate behavioural science across government and throughout civil society. The report *EAST: Four Simple Ways to Apply Behavioural Insights* has become a

flagship for the BIT. Described in Halpern (2015) as the four "central approaches" (p. 10), EAST is a heuristic of heuristics, an intended nudge for the nudgers that provides a mental shortcut for effective communication. This mnemonic, standing for Easy, Attractive, Social and Timely, not only typifies the BIT's policy recommendations, but shows the team moving towards a more educative and training role and away from making policy itself. A second trend is that the BIT is being employed to use its expertise in running Random Controlled Trials (RCTs). *Growth Vouchers* is typical of this, for which the BIT led an evaluation of whether advice given to small businesses improved business performance. While the BIT did draft communication between the government and businesses, the purpose was to gauge the effect of the vouchers themselves, so no behavioural insights were used or gained. As part of its increasingly educative purpose, the BIT produced a report (*Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials*) to inform trials ran in the public sector.

The BIT claimed that this change of course is a mark of success, but other explanations are more parsimonious. The first one is that the BIT ran out of nudges to recommend to central government, and so moved on to other projects. The second is that the BIT is more concerned with making existing policies and institutions more effective than providing innovative policy recommendations that may make a case for reducing the role and size of those institutions, as the original intent of the Conservative Party, and the promise of nudge theory, was.

5.3 Policy Recommendations

Below is a discussion of the BIT's published policy reports, describing the policy recommendations they contain. Most discussed were trialled before the BIT's review in 2012, even those published after that date.

Included are some policy ideas that were trialled without positive results. Cumulatively they show that the BIT has overwhelmingly tried to re-frame the way people see issues or win attention by altering the manner of government communication and the way information is presented, as sceptics in government alleged.

5.3.1 Basic Approaches

EAST and MINDSPACE are both mneomics intended to guide government communication, and their influence is seen throughout the BIT's policies. Twenty-nine of the forty-three policies are exclusively or predominantly a change in the method of communication or provision of information. The methods of communication can be categorised as:

1. Personalisation: Changing communication to bring abstract concepts closer to subject.

2. Aesthetics: Making communication more attractive.

3. Simplicity: Effectiveness of communication is increased by making the message easier to understand.

4. Timeliness: Communicating a message at a time most likely to inspire behaviour change.

5. Saliency: Increasing the visibility of some aspects of the message. This usually takes advantage of attentional bias and anchoring.

6. Selecting a messenger: Uses biases towards authority to promote a message.

These are all intended to propel the recipient them towards a behaviour, but provision of information usually is not, as it appeals primarily to

system 2. Prospective nudgees are "significantly more likely to register stimuli that are novel, accessible and simple" (*Better choices, better deals*, p. 21), which provides a cost effective way of changing behaviour. Attempts to promote behaviour change using information was tried under New Labour. From traffic light symbols on food packaging to attempts to reduce carbon dioxide emissions, these policies were seen to have failed because they "relied wholly on providing information" (Young & Middlemiss, 2012, p. 744). The BIT's approach is to strategically re-frame choices by selectively choosing what information to present and how.

At least eight (more depending on how reciprocity and other effects are regarded) policies are based on using social pressure. Applying social pressure, appealing to herd instinct or what the BIT euphemistically call 'the power of the crowd', was generally effective. Social pressure is well established in behavioural science, and TS's section on it makes a compelling case for its effectiveness. Engendering reciprocity creates a similar affect, but rather than creating a circumstance in which people imitate behaviours, it creates a commitment to social norms.

Another nudge mechanism is the use of defaults. This narrows the choice architecture facing someone from a multiplicity of options, to the choices desired by the nudger, or to make an opt-in system to an opt-out one. This was recommended on just two occasions by the BIT, but there were other times that a heuristic was provided. Removing friction and changing incentives is central to the BIT's approach and underpins many of the recommendations they made.

5.3.2 The Early Period (Policy Publications 1-5)

During the early period the BIT took policies straight from *Nudge*. In *Better choices: better deals* the BIT offers information as a solution to UK consumers having a low rate of price comparison in relation to European counterparts (1). The BIT does not attribute this to inertia, but that is clearly the only the bias this action could be aimed at addressing. The BIT advocated 'Mydata', a programme intended to allow consumers access to information that companies hold about them, adapted from a similar policy in *Nudge* (TS, p. 173). In 2015 the BIT expressed disappointment with the progress of 'midata' (which had apparently undergone a name change in the meantime), seeing only 1.8% of accounts switch supplier. In the same report it was proposed by the BIT that the credit card industry voluntarily make available annual statements that not only provide comprehensive information of spending habits and product information, but also facilitate comparison and ultimately switching between service providers (2), also taken from the pages of Nudge (TS, p. 143).

Displaying important information in a more eye-catching place is a favourite technique of the BIT. An early example of combining these techniques was the redesigned presentation of Energy Performance Certificates (EPCs) to become more attractive, and ensure the information most likely to change behaviour is placed in the most salient position, as well as displayed more simply (8). In this instance the BIT also sought to remove friction by providing information about small changes that the subject can make to reduce carbon emissions. An early experiment with trusted messengers as well as of timely intervention involved recruiting the Royal Institution of Chartered Surveyors (RICS) as a medium to recommend loft insulation as people were moving

homes when friction costs are smaller. Friction was minimised for people not moving house by offering loft clearance along with loft insulation (5), apparently increasing uptake threefold (Halpern, 2015, p. 3). Prompting behaviour at timely moments was trialled again, by asking people to donate when writing their wills. This increased the number of people donating in their wills from 4.9% to 10.8%, which rose to 15.4% when combined with a social norm (22).

One area of policy where the BIT's impact is most felt is the manner of letters sent by the government. The days of receiving a benign government letter not preceded by an aggressive instruction, such as 'ACT NOW' on the envelope demanding some action are a quaint memory. This was pioneered in the early period of the BIT by experimenting with tax collection letters, and was noted by Halpern as the policy that convinced bureaucrats of his approach (Halpern, 2015). This involved personalising letters (13), simplifying the content of letters (11 & 13), appealing to social norms (10 & 11), engaging the subject's personal beliefs about tax, reciprocity and commitment (15 & 17), as well as what can only be described as threats (12 & 16). Personalisation, appealing to egocentrism, seems to have been the most effective, and became important to the BIT's approach since then. Applying social pressure was the least effective, it was shown that appealing to the social morality of dentists and doctors had no discernible impact on this occasion. The success of these letters was enough to convince leading civil servants that government should now be concerned with creating an affect in their letters instead of just conveying the information once thought necessary, whether that affect is shame, fear or appealing to vanity.

The BIT parlayed this success into using the same methods to increase charity donations. Personalisation increased participation in a payroll giving scheme from 5% to 12%, and then to 17% when combined with a reciprocity device. Social pressure also proved effective. A trial amongst HMRC employees saw two e-cards sent out during winter (20). One had written messages from colleagues who were current donors, which had a success rate of 2.9%, while the same message accompanied by a picture of the donor had a success rate of 6.4%, an outcome described as "striking" by the BIT (p. 18). The handing out of sweets increased charitable giving by increasing feelings of reciprocity (21). Interestingly, a follow-up trial revealed that people who received sweets twice were less likely to donate than (9.7% vs 5.8%). The BIT also offered feedback on energy consumption, compared to social norms. When the consumer was relatively energy efficient, a smiley face was given to prevent a 'boomerang effect' of a reversal of desired behaviour towards the mean (7).

The BIT also tried to create 'collaborative consumption' in the hope that groups of consumers will self-organise to combine their purchasing power to make better deals, or to share or trade goods with each other. This was to be achieved by providing a prize fund, of up to £30,000, for the most innovative scheme. Providing a financial incentive for taking up energy efficient products that is positively affected by the number of people in a community that opt into the offer is apparently an example of social pressure because it encourages neighbours to make their own decisions transparent and cajole others to match their decision. A ten percent discount for two households, fifteen percent for three, and twenty-five percent for five households was offered (4).

In addition to using communication methods, the BIT tried changing defaults during the early period. A trial was run on private sector employees participating in a payroll giving scheme to see if a default opt-in for automatic escalation of donations, at 3 percent annually, would increase the amount of donors who opt for automatic escalation (19). The result was an increase from three percent to forty-nine percent. In an effort to show how charitable giving can be increased the BIT sent out three different letter frames to infrequent donors. Neither of the three letters differed much in rates of increased donations, all of them receiving about 3% positive responses, but the total amount raised was more than double when options for the donation amount were given £2, £4, £6, £8 and £10 instead of £1, £2, £3, £5 and £10, while the latter option increased the total amount donated over the letter that offered no specific options (18). Nowhere did they acknowledge concerns about parting with money by psychological manipulation, although this second use of defaults was a very soft default, as much a provision of heuristic as a change of default.

A notable feature of this early period is that it is the only time that the BIT offers an explanation of what bias it is responding to, but by the third policy paper this is no longer the case. That paper, which featured the series of letters sent to tax cheats features recommendations that do not seem to derive from any bias at all. Even policies directed entirely at the reflective system can have cognitive biases attributed to them where none are necessarily present. The various financial incentives offered in the Green Deal (3) were said to "tap into people's aversion to anticipated regret" (p. 11).

5.3.3 The Later Period (Policy Publications 6-22)

An example of the provision of information failing to produce behaviour change was a trial run with department store John Lewis, imitating one performed in Norway, that tested whether altering energy labels on washers and dryers to include the lifetime running costs changed purchasing choices (41). In this instance only washer-dryers saw a significant positive effect. Altering the style of communication has proven to be more effective than simply providing information. Personalisation was found to be effective if addressed to ethnic groups as well as individuals. By changing the message received by applicants to become police officers to one cached in terms of benefits to themselves and their community, the BIT reduced the rate of failure by BME applicants at a troublesome stage of the application process (37).

Reducing the provision of information can also achieve an outcome. It was found that school students responded more to hearing the lifestyle benefits of university than career benefits (43), showing how the myopia of students could be used to their benefit. Appealing to authority bias was not effective in the latter period. An attempt to use trusted messengers to advise people how to more effectively use their boilers and thermostats had a visibly negative effect (42). The BIT continued to experiment with different forms of communication, trialling what modes of communication are most effective to make employers rethink their decisions to employ migrants without leave to remain (38).

Often the old policies, especially the ones that use communication, were re-hashed for new audiences and purposes (28, 29, 30, 34, 35 & 39). Another example of choosing a salient moment to communicate a message that was not successful was an attempt to promote smoking

cessation to newly pregnant women. This failure prompted the BIT to note that it is "important to test and trial interventions before rolling them out" (BIT, 2015b, p. 15), and is reflective of failures to reach lower socio-economic groups.

A conspicuous policy was the redesign of the government's organ donation website (23). The BIT trialled eight different webpages designed to increase organ donation, each with a different message communicated. All but one of the seven designs performed better than the base, which simply stated "THANK YOU: Please join the NHS Organ Donation Register". The webpage that underperformed featured a picture of a crowd of people, while the BIT was surprised that any of the seven alternatives performed worse than the base. The two pages that performed best were the one that framed failure to donate as a loss ("three people die every day because there are not enough organ donors"), and one that emphasised the reciprocal nature of organ donation ("If you needed an organ transplant would you have one?"). One conclusion from this trial is that the aesthetics of the website was not significant, it was the incentives and social norms, the message itself, that motivated behaviour change, as the page that simply changed visuals decreased the success rate. Something similar was done at a later date, during which it was found that a carousel decreased the efficacy of webpages (34). Therefore there are two examples of the BIT trying to make something more attractive, with the final effect opposite of what was expected.

Another insight from the BIT led to the reforming of the pupil premium from simply giving schools more funding depending on how many disadvantaged children they are responsible for educating, to one that distributes awards for improving the performance of disadvantaged

pupils (36). This changing of incentives is the result of no bias and uses no bias, strictly engages system 2, and shows the BIT converging on 'normal' government policies, actually just tweaking existing ones. One policy that the BIT advocated but was not adopted was lotteries with prizes to reward people who register to vote (Halpern, 2015). These policies appeal primarily to system two and constitute conventional government intervention, albeit more targeted and evidence-based. They impose costs on taxpayers (though perhaps for long-term savings) and show behavioural science not being used to shrink the level of state but to extend its reach.

The end of the sample shows the BIT increasingly working on changing the behaviour of government employees, often with heuristics to help them handle complicated tasks. This has been done for medical prescriptions to reduce the need for handwriting and resulting misunderstanding (31), and for social workers to improve their decision-making (24, 26 & 27). The BIT stopped policy recommendations altogether, or sold replicas of policies to different countries. Eleven papers from this period (compared to one from the early period) feature no policy recommendations, and instead advise organisations about integrating psychology into their work, in a manner reminiscent of the 'think pieces' of the PMSU. The educative function the BIT has taken up involves educating people about their own cognitive biases, as well as educating NGOs about the ways biases affect their services. Frontline policies that have focussed on educating the general public have been rare for the BIT. One such paper was based on West Sussex, intended to to improve the parenting skills of foster parents, specifically their resilience, well-being, engagement with the community, and stress-handling ability..

5.4 Summary

The BIT's eventual transition from a government body to a charity signalled a decrease in emphasis from giving advice to government (though it allowed the BIT to give the same advice to governments around the world and get paid for it), to one that primarily focusses on educating organisations about behaviouralism. The chronology of the BIT's recommendations is revealing. Seventeen of the forty-three published policies were published before the end of 2012, but this understates the quantity of work done before that date. The first real policy report was not published until several months after the BIT was founded, and so some the work for some reports published in 2013 onwards were also have been done before the end of 2012. Some of the policies in the 2013-2015 report, which contains fourteen of the recommendations of the latter period, are conclusions from trails ongoing during earlier publications. Furthermore, that report also contains policy experiments that failed, while earlier policy trials that failed are not published, but are discussed in other sources.

That some policies in the early period are taken straight form *Nudge* is not incidental. Halpern (2015) stated openly that the BIT looked for the "most prominent and best-evidenced ideas from the wider literature" to "provide some early quick wins, and help establish the approach" (p. 52). These early recommendations seem more deviceful and diverse than later policies. Defaults were used only in this period, social pressure was used in creative ways, communication methods were experimented with and policies directly affected people's lives. Repetition is inevitable, but in the latter period there is movement away from policies that directly increase welfare of people, and towards administrative efficiency. Even in the early period, however, this is in evidence as shown by the tax

collection letters, and given the proximity and mutual dependence of the charity sector and government, in efforts to increase charitable giving as well.

The BIT not only seems to now be more focussed on generally educating governance organisations about behavioural science than producing policies, but throughout its existence the BIT has been often reduced to the role of cheerleader for developments taking place in commerce. This is evident in the support for Opower, efforts of mobile phone manufacturers to disincentivise mobile phone theft, as well as encouraging e-cigarettes. One possible reason for this is that the number of possible nudges is finite. The dismissal of the BIT that it had little to contribute because it was really just about changing the style of communication has gained support from its record.

6. <u>Comparison and Analysis</u>

Of the forty-three policy recommendations, twelve fully accord with nudge theory, though another nine practically do, and further ten comply with the moral criteria of the theory. Some of those that do or practically do appeal primarily to rationality or tries to overcome a bias rather than use it, suggesting no great contribution of nudge theory to what is already an unremarkable set of achievements. In addition to administrative and ideological continuity, a possible explanation is general difficulty translating theory and evidence into policy. This is often hampered by political factors not relating to the theory and research themselves. With most evidence-based policies one can see cultural and political rationalities triumphing over technical rationalities (Lin, 2003). Halpern discusses political obstacles throughout '*Inside the Nudge Unit*',

mainly in the shape of scepticism from civil servants. Some of this was also cultural, relating to ethical concerns about the use of psychological techniques. Further, the risk of upsetting the working of the government machine deterred bureaucrats from allowing the BIT to trial policy ideas (Halpern, 2015). Other potential obstacles such as financial considerations are not likely, as the BIT was considered a cost-saving exercise.

Despite some of the early policies being taken straight from Nudge, there is no relationship between the chronology of policies and their accordance with nudge theory. Indeed, the examples that TS provide are themselves in constant tension with TS's criteria. They identify a necessity to be insistent with their moral criteria precisely because of its vulnerability. Two things determined whether a policy was compliant with nudge theory. The first was that the policy focussed on changing the behaviour of government employees. The second used mainly information to achieve outcomes. The reason for the first determinant is obvious, the moral risks of nudging someone who has already entered into a voluntary contract as an employee are far smaller. For the second determinant, the reason is that information is less burdensome and will rarely violate the TS's moral criteria. However, information could be classified as 'fuzzy nudge' as it relies mainly on system 2, and gives pause about nudge theory's utility.

The provision of information is usually just considered a normal policy to correct a market failure. Even the policy of 'MyData', from the first year of the BIT'S existence, was already being worked on by the government before the BIT began to explain the need for it in terms of cognitive biases. 'Traffic light', actually guide behaviour to a far greater extent than the BIT's information-based output, shows policy movement towards

behaviour change without formal psychological impetus, and altering the presentation of information to achieve it. The reasoning for introducing traffic light labelling onto food packaging followed much the same thinking as the BIT's information-based programmes, as they aimed to preserve consumer choice while hoping to increase some choices and correct a market failure, reflective of classical economic thinking, not behavioural economics. While Halpern (2015) claims this nudge-type policy was a success, and it is one of his favourite examples of such a policy, the bulk of evidence suggests that improvements in diet may have come from other sources (Loewinstein, Asch, Friedman, Melichar & Volpp, 2012), perhaps cultural changes that precipitated the move towards government action in the first instance.

6.1 Mechanistic Criteria

6.1.1 Cognitive Biases

The first step to a nudge is to identify a problem. There would be a misstep if the BIT problematised an issue where no cognitive bias is in tension with the desires of the inner econ, even if the eventual policy better fulfilled those desires. The result of this would be a normal government policy rather than correcting a cognitive bias. Policy recommendations overwhelmingly do aim to alter decision-making that can plausibly be conceived of as hindered by cognitive biases. Moreover, all of the biases described in *Nudge* are the basis of an action by the BIT. For example, providing a predictive model for social workers is an attempt to temper overconfidence and optimism.

Perhaps predictably given the preponderance of communication, the idea of framing issues differently permeates much of what the BIT has done. The classic example is to use people's bias towards loss aversion to frame a possible gain as a possible loss. The BIT did this, but did not frame it as a loss to the the subject in question, but instead to society, meaning that it often overlaps with using social norms. During tax collections, for example, the HMRC made the social costs of unpaid taxes known.

The problem with framing is that it is is ubiquitous, no one is exempt from constructing their own frame, or world-view. For a re-frame to be considered a nudge, it would have to be shown that or that it presents a world-view closer to reality, or at least that the subject prefers the new frame to their old one, because otherwise the initial frame could not be considered a bias, any more than the government's frame could be. In some cases, such as headlining letters with non-specific threats such as "pay now" in red ink, that is doubtful. Using social pressure to promote charitable giving is effective, but it is unclear what frame is being shaped here (and it can only be a response to apophenia or status quo bias). The information may be accurate (but incomplete), the frame being shaped is potentially one that instills shame and self-doubt, which, if the frame is not an accurate reflection of the subjects' character, effectively imposes a new bias which leaves the subject less happy by distorting their self-image.

It is difficult to distill the biases present in most problematisations. Everything conceived of as a policy problem related to behaviour can be seen as a framing problem, and creates a need in the view of policy-makers to shift from one mental structure to another. Furthermore, inertia is present as a bias in almost all of the issues the BIT sought to

change. Any inaction by a subject that the BIT believes the subject may wish to change when thinking as an econ or deliberative self would be, in the view of the BIT, inertia. The sheer number and mundaneness of the potential 'nudges' possible using inertia and framing as a basis demonstrates the need for a stricter moral criteria to justify nudges, without it policy-makers can arbitrarily impose a new reality on people.

6.1.2 Behaviour Change

Using Baldwin's (2013) classification, twenty policies use re-framing techniques, sixteen appeal exclusively to system 2, while just seven alter choice architecture. This understates the extent of the BIT's reliance on the reflective system, however, as even the re-framing techniques rely to a considerable extent on the subject engaging this system. All information has a re-framing effect as well, as it has to be decided what information should be presented, and this can be used to create affect. Halpern (2015) describes the BIT as operating within a new standard of truth, one where what is communicated has only to be true in itself. What does not matter to the BIT is providing citizens with comprehensive information, that may allow them to make a decision that accords with their reflective system, as even TS concede. What is communicated is that which is most likely to change behaviour, with the effect that it could be misleading. An example of this is advertising the apparent social benefits of paying taxes, which is arguably untrue, or at least misleading if a full assessment of taxes and government expenditure were to take place. In this way, the provision of information ceases to be "highly libertarian", and contrasts with the historic role of government in the provision of information, although in most cases the BIT's information has been too ineffective in changing behaviour to be alarming.

Presenting information can be highly libertarian if the information is correct and comprehensive. While providing information does potentially correct a cognitive bias, it still engages the reflective system and does not necessarily shape choices. The BIT's experience with the 'MyData' scheme, the traffic light scheme, labelling on white goods and the failure to encourage smoking cessation among the newly-pregnant perhaps indicates that a simply informative approach has limited behaviour change potential, certainly compared to the generally more successful communicative methods. This shows that the libertarian, or moral criteria, of nudging is in tension with achieving policy ends.

The dependence on engagement with the reflective system for policy ends gives reason to doubt the utility of nudge approaches. TS (2008) make a point of explaining that a lack of information does not necessarily mean that people are not econs, as that requires that people be wrong in a systematic and predictable way (p. 7). So while information can be considered a nudge of a sort, the great number of such policy recommendations in the BIT's output make a mockery of the idea of psychology in policy. If this is to be the approach, it is difficult to see the need for a BIT at all.

The educative function of the BIT is even more anomalous. In seeking to challenge and confront biases the BIT is acting in a way that writers have seen as the antithesis of nudging, and converging upon the approach proposed by some of the critics of nudge theory. These are actions that take people away from their automatic systems and to begin engagement with their reflexive, or 'think' systems and shows the BIT acting in a manner that critics of nudge theory recommend as an alternative.

The use of incentives, especially financial incentives, does not accord with nudge theory, as they provide rational reasons for an individual to change their behaviour. As TS (2008) put it "a nudge is any factor that significantly alters the behaviour of Humans, even though it would be ignored by Econs. Econs respond primarily to incentives" (p. 8). Incentives therefore depend on people's reflexive systems, and fit the conventional model of economics and policy-making. These policies contradict nudge theory, but it seems fitting that the idea came from the BIT because it fits the BIT's mandate of providing cost effective and innovative (however superficial a lottery and financial reward may seem) solutions to problematised issues.

6.2 Ethical Criteria

The BIT does not promulgate moral restraints in any of its publications, but there is one instance where they moderate their recommendations for something more appropriate. Halpern (2015) dismisses many of the moral concerns people have about psychology being used by policy-makers, and defends usual democratic restraints as sufficient. Concerning the American libertarian line of criticism, he claims that fears were actually of a "more effective administrator" (p. 42), and not psychological manipulation. He pays very little attention to the leftist critiques, perhaps because their main thrust is the political intention of the Conservatives, and he is unconcerned with defending the Conservative Party.

6.2.1 The BIT and the Inner Econ

While the desires of the inner econ may elusive, it can be said with certainty that objections could be raised to every policy. Charitable giving, at least giving to specific charities, will always have objectors, so will the payment of taxes, and libertarians consider voting to be an act of violence. The only way that it is possible for government actors to achieve this is to assume that it is permissible to ignore anomalous concerns. Even within this criteria, the inner econs of large numbers of people are violated several times. Leaving aside the issues of personal privacy 'MyData' presents that will inevitably trouble some people, even a seemingly benign policy recommendation such as a price comparison website may not accord with the inner econ or the deliberative self. Internet use has a considerable carbon footprint, especially cloud computing as 'MyData' uses on an enormous scale, and this may not be cost effective for the subject, of society generally, or even of government policy commitments (Walsh, 2013).

Another example is the lottery scheme to increase voter registration, completely harmless on the surface. Given the UK's first-past-the-post district-based electoral system, swinging the vote is a 'Black Swan' event, leaving British voters more likely to suffer a fatal injury while travelling to their local polling station than they are to cast the decisive vote (Pagels, 2014). Nudging people towards voting cannot be said to benefit the subject, even if the subject considered the choice of candidates offered to be as good as they believe their death undesirable. Voting passes the new year's resolution test, in the sense that people will resolve to vote but rarely to not vote, but does not accord with the inner econ, which only shows the bias of that test. It relies on a collectivist logic, not on doing what is good for individuals, but this is more a sign of how nudge theory does not live up to its claims to libertarianism, and inconsistency within nudge theory, than how much

the BIT deviates from nudge theory. It is likely that in this case, it is not the collective or the body politic that gains from increased voter turnout, but only politicians themselves who attain undeserved legitimacy from the process.

The preponderance of communication in the BIT's policy recommendations is not only un-innovative but may be disagreeable for people subject to them. They can be irritating and patronising, but more importantly they also represent a lowering of standards of honesty in government. They can impose biases, as well as be threatening, causing fear, or shame-inducing. Social pressure is even more troublesome, evidently a powerful tool that can take hold and shape behaviour despite persistent negative attitudes towards that behaviour on the part of the subject (Young & Middlemiss, 2012). This means that it causes people to adopt behaviours that they regret, effectively rendering the policies recommended by the BIT as bullying. In the case of using social pressure to increase charitable giving, it can be assumed that some donors would have otherwise preferred to not donate, seeing the charity sector as self-serving, ideological or comprised of failures or would have preferred to have donated to a different charity. Additionally, it is known that publics are particularly wary of nudges that cost them money, which all the efforts to increase charitable donations do.

By using social pressure and narrowly framing some issues, the BIT is not nudging people towards rational decisions. They are in fact introducing irrationality to a rational situation. The overall impression is that the BIT is concerned with creating more effective administrators, striving for current policy aims, and that helping people to fulfil their own preferences is not a consideration. This ideal is mentioned a few times, most notably in *MINDSPACE*, when *Nudge* was most prominent in the

thinking of the BIT, and in later general papers when the approach was being explained and sold to new audiences. However, this appears to have been salesmanship, and it has not bled through to policy recommendations.

6.2.2 Choice Restriction

Given the difficulty of providing recommendations that accord with the inner econ, providing recommendations that are easy to avoid becomes more important. None of the BIT's policy recommendations are coercive, and so all behaviours the BIT is aiming to achieve can be avoided. On several occasions, the BIT also showed disregard for the freedom of businesses and taxpayers, but this is mirrored by TS, even if it remains a contradictory position, and there is concern expressed that the incentives they change disincentivise some behaviours. The criteria in nudge theory may refer only to the person subject to a nudge.

The chief concern is that most policy recommendations are about methods of communication and increasing information, and so the attempt to re-frame a subjects' thoughts are completely unavoidable once the text has been communicated. Communication and information do not so much structure choice as they provide the government with an unavoidable influence over the mind of the subject. In promoting the utility of taxes, the BIT is reinforcing political norms that benefit some individuals and groups at the expense of others.

The mechanisms most likely to create behaviour change were shown in the seven policies that directly altered choice architecture. Most

prominent was the automatic escalation in charitable giving that increases enrollment from three to forty-nine percent. This was probably the most significant change of all the policies. By comparison, the prompting of donations in wills only increased donations from just under five percent to less than sixteen percent at its highest point. The BIT explicitly stated that it thought it was inappropriate to set defaults in the latter example, but it is difficult to see what distinguishes it from automatic escalation.

This reveals a paradox in nudge theory, for to show that a nudge is avoidable, it needs to be avoided. Nudge theory's success therefore depends upon its failure, and this degree of success may be too great to show that this particular recommendation was easy enough to avoid. For a theory to be built on people's mental frailties, and knowing that the recommendation does not accord with many people's preferences, it is certain that some of the increase in the case of automatic escalation was the result of exploited subjects; unfortunately the BIT did not publish data showing how many recanted when they realised what they had agreed to. The potency of setting defaults, what the British people rebelled against in the case of organ donation (Halpern, 2015, p. 52), is the reason there are so few examples of it. Further, it exposes the threefold protection nudge theory offers as a tightrope that policy-makers have to balance upon, and the tension between effectiveness and TS's ethical criteria.

6.2.3 Rawlsian Principle and Transparency

Halpern (2015) echoes Sunstein in insisting on transparency for all the workings of the BIT. He does not specify whether he means that they must be published in some capacity, or discussed in a public forum or some other criteria. Halpern knows that inertia means his organisation

will not necessarily be subject to scrutiny. He also justifies the BIT's freedom to work upon the legitimacy of elected policy-makers who have the final say on all decisions, although in most cases most of the work was approved by unelected civil servants. Nevertheless, many things are done under a cloak of democracy and Halpern is happy for this very sensitive and potentially dangerous set of policy tools to join them. The only possible way of judging whether a policy is transparent is whether the person subject to them is capable of realising that the choice architecture, information or communication they are faced with is trying to change their behaviour. Even with this as the criteria, it has to be determined how attainable this awareness has to be, and by whom. It can be safely presumed that the most vulnerable people are those least likely to be aware of this fact.

The problem with the sample is revealed most by the question of transparency. All policy recommendations in the sample are transparent to the extent that they are published, but not all policies have been published. Certainly all of the policies published were thought to be publicly justifiable. One example of an unpublished policy that provoked protest is one for jobseekers, who were forced into "wasting time on mumbo-jumbo personality tests" (BBC, 2013). This was not included in the sample because it was not published. Merely publishing policies, even if what is published is not biased, is clearly insufficient for transparency, and the Rawlsian principle even more so.

Sunstein was correct to improve upon the Rawlsian principle by adding the requirement for transparency. However, it is untrue that the BIT has been unaware of the need to restrain its activities or misunderstood why employing social psychology in policy is potentially dangerous, the reason they opted against defaults in some circumstances is because

they believed it be an immoderate option. However, in most cases it has been civil servants and ministers who have prevented the BIT from conducting trials considered unethical, although the BIT even bypassed normal procedure sometimes (Halpern, 2015, p. 5).

6.4 Efficacy

The BIT's limited achievements and diminishing influence, in addition to the little use it found for nudge theory itself, shows the limitations of nudge theory. Even in areas where the BIT has been effective, there are reasons to doubt whether this success will endure that emerge from the BIT's own publications. The first reason is that gains made by communication are the result of novelty, something the BIT acknowledges. The second is that the BIT's mode of communication makes government communication resemble that of commercial interests. The BIT speculated when reporting on website designs that the reason one design had been a deterrent rather than an attraction was that it had a commercial appearance. Yet all of the communication techniques used by the BIT follow techniques that are widespread outside of government and academia, so not only will the novelty be short-lived but may provoke public cynicism.

The informative approach has not been successful. For example, 'myData' eases friction, but may still be too difficult to use, and still requires activity and motivation, so does not tackle the base problem of inertia. The attempt to target pregnant mothers with anti-smoking messages was not successful either, much like previous anti-smoking campaigns. This is one example of the BIT targeting a particular socio-economic group and failing, supporting the evidence of the literature that some populations will be more difficult to nudge. Psychology is in the midst of a replicability crisis, and the BIT is not

immune to that. In their organ donation website design they were surprised that their expectations about what would be successful were confounded. Many times they appear to be trying different ideas, and are unable to explain why some fail and some succeed. Psychological policies may be more context-specific than TS hoped. The organ donation website policy shows the BIT at its most directionless and unmoored in theory, but it is not wholly atypical.

The conflict between the automatic and reflective system is yet to be resolved. Halpern (2015) implicitly supports the critics of nudge theory when he states that conversion, not compliance is needed (p. 274), and that it is better that we learn about our biases. In this instance the BIT seems to be adopting the educator role that Martin (2015) stated his preference for above the BIT. Requiring reflection to change behaviour does mean that some of the hypothesised effects of nudging will be lost, such as the decrease in cognitive specialisation, causing more decision fatigue, a reason TS give for sub-optimal decision making. Halpern provides another reason to doubt the long term efficacy of nudge when making the claim that well-being is negatively affected when people think that their agency is being violated (p. 260-1). Some nudges, such as providing information, supports agency, but most nudges and the technocratic nature of the BIT do not.

6.5 Summary

Observing the issues problematised and acted upon by the BIT, it is clear that the BIT has been restricted to valence issues, reflecting the thinking of higher-ups. This means that policy recommendations meet Halpern's moral criteria, and while falling short of TS's inner econ, reflects the current moral consensus, at least of that in government. In hindsight, it was probably to be expected psychological policies would

be measured against existing norms in government by those in government, and not against TS's libertarian-infused ideas. There are no policies that cannot be justified under present political discourse, but a few would be controversial, the automatic escalation of charitable giving most of all. This violates all three of TS's moral criteria, the only published policy that does so, and involves a mechanism, changing defaults, and an end, involving the loss of money, that people express apprehension about.

It should also be no surprise that a theory that strives to satisfy a libertarian ethos would not provide many tools for policy-makers. The reliance on information and communication, and even more so the reversion to policies informed by classical economics, as methods of behaviour change shows the difficulty in creating psychological policies that protect freedom of choice. An even more significant reason to doubt the impact of nudge theory is not just its lack of policy tools, but the abundance of other explanations for the BIT's policies.

To isolate the impact of nudge theory on the policies of the BIT would require to eliminate all competing explanations. Yet *Nudge* and Halpern (2015) are replete with examples of nudge-type policies implemented prior to nudge theory, sometimes by centuries. Other policies, such as sending text messages to remind people of court appointments, were already being used on the initiative of service providers. The provision of information and communication are normal government policies. Further, all the techniques are normal communication techniques: salient information is usually placed in more visible positions; language has historically been as simple as was previously thought possible; the aesthetics of communication has always been a consideration. It is difficult to even think of the BIT as agents of change in this instance,

because the possibilities of all their policy introductions must have already been known. What changed was the extent to which these techniques were embraced. This also undermines the claims that behavioural science has to being able to impact policy. It has always been used by policy-makers, on an ad hoc basis, and has never been a great driver of change. Using behavioural science within TS's boundaries will only curtail its utility further.

7. <u>Discussion</u>

Policy-makers have always used psychology to change the behaviour of people in provisional ways. Nudge theory, drawing on an ever growing body of knowledge in psychology and behavioural science, is the first attempt to create a framework for using social psychology in policy, while the BIT was established as the first attempt to systematically make policies using psychology and behavioural science. This had been done tentatively before in the UK, with a direct link from those attempts to the BIT, but it was unpopular and did not last long. Nudge theory, an exciting, novel and fashionable concept provided new impetus, promising that freedom can be preserved, and the state reversed, in conjunction with welfare gains.

The definition of nudge theory in this paper is the one supported by the description of that theory by TS. However, even their examples, including those imitated by the BIT, do not match that description. The issue is that the restrictions within nudge theory, though they enable it to be supported across the political spectrum, narrow the use of behavioural science in policy to fruitlessness. .

For his part, Halpern (2015) is unconcerned whether or not what the BIT does accords with nudge theory, though in describing almost everything the BIT does as a nudge gives the impression that he thinks nudging is whatever the BIT does. The discord between nudge theory and the BIT could be because of apathy on the part of the BIT, the predominance of extant policy rationalities and aims, or possibly because of the inapplicability of nudge theory itself. The evidence suggests the latter, in part because there was an initial desire to use nudge theory, use of its motifs and even wording, and in part because the tensions in application reflect tensions within nudge theory.

Many policies do not even draw on behavioural science, and rely on conventional economic thought instead. John's (2013) suggestion that the BIT is in fact a 'skunkworks' unit, an institution within a larger organisation charged with finding solutions that other units, when immersed in their normal tasks, would overlook or not be aware of, could be an appropriate model. This model suggests that the BIT has become purely pragmatic in its approach, and that in its competing pragmatisms is finding decreasing use for behavioural science.

There is no reason to quibble over the disapplication of the mechanistic criteria, but the move away from that does show the limitations of nudge theory. Combined with the move away from giving policy advice altogether, as well as the early dependence on policies taken directly from *Nudge*, it is clear that nudge theory is a finite policy tool. One finding does suggest the effectiveness of nudge: changing defaults was very effective, while what TS describe as "highly libertarian" nudges were considerably less so. The policies that used a cognitive bias in a positive way rather than engaging with the reflective system were more successful, validating the logic behind nudge theory at least.

More problematic is the BIT's disregarding of the moral clauses of nudge theory. Most egregious is the willingness to nudge people towards behaviours not in accordance to their inner econ, or irrelevant to people's health, wealth and happiness. The BIT seeks to nudge people, at least when it did, towards behaviours that the government wants people to have. This represents a complete derogation of the libertarian sensibilities that should underpin nudging. This places the BIT alongside the preponderance of other government actors, and it's recommendations alongside most other policies in helping the government achieving administrative aims rather than permitting or enabling citizens to achieve their aims. In this interpretation the BIT derives its legitimacy from being ultimately responsible to democratic government. Nudge theory, however, is sensitive to the fact that using psychological techniques to change behaviour requires more than ordinary answerability.

"Highly libertarian" policy recommendations preponderate the BIT's output, and the BIT has admitted to shying away from more aggressive policies, such as changing defaults, that seek to bypass system 2. Even just re-framing of information is an exercise of power that violates most definitions of libertarian, though all behaviours the BIT has attempted to change remained voluntary for the nudgee. TS's ideas of the inner econ and the new year's test leave arbitrary power in the hands of policy-makers, and is an impossible standard for them to achieve. Sunstein's claim that nudging needs to be transparent is shown to be particularly difficult to pin down, there are no instructions provided by TS. While the BIT refer to TS's other moral criteria in their policy papers, transparency receives no attention, and represents a hole within nudge theory.

Libertarian paternalism is shown to be a contradiction, not a paradox. The libertarian aspect in nudge theory may even be, given the tensions within it, and TS's own examples, a selling tactic, and not something TS take too seriously themselves. TS are conventional American liberals with sympathy for elements of libertarianism that have a strong established empirical support, so their framework may just be intended as a general guide. Their description is wholly more benign and libertarian than the examples they provide; the BIT shies away from their examples but falls short their description. A cynical interpretation is that nudging was an attempt to win conservative support for interventionist policies. In any event, there was only a fleeting chance that the BIT might be infused with a libertarian ethic. Governments essentially do two things: they forbid people from doing what they want to do, or they compel people to do things they do not want to do. To begin enabling people to do what they do want to do, by removing their own internal impediments to that, could have been revolutionary in governance. The BIT did not, however, set out to do that.

Psychology in policy has been justified by the left in terms of market failure, but commercial products that use and overcome cognitive biases are far outstripping what policy-makers can do. There is neither a technique, nor any end that someone may want to be nudged towards that could not be satisfied by the market, and the market is already moving in that direction. Just as governments have lagged behind businesses in using psychological techniques, so too are they inevitably less dynamic in meeting the wants of a population that wants to purchase nudge services. In moving away from working only for government, the BIT has essentially conceded this point itself. The BIT's staid system (when compared to the market) of RCTs will not generate

insights as significant or as quickly as the market will. This should not be surprising, for if the demand for nudges were satisfied by the free market, then it will be satisfied in the same way that most other demand for health, wealth and happiness is. *Nudge* and the writings of Kahneman have already done more to advance awareness of cognitive biases than the BIT can feasibly do. Cultural trends and marketplace for ideas are pushing people towards more healthful options, and "strong evidence that consumers are sampling the growing number of nudges offered by markets", such as stickk.com and numerous weight loss and addiction apps (Marlow, 2014, p. 26).

Halpern (2015) accepts the limitations of government intervention in a couple of instances. The first is in his expressed hope that the BIT's policies would have a snowball effect. The 1.8% of mydata users who switched may represent a very small number, and a failure of the policy, but the BIT hopes that culture is changed by it, information is disseminated, and social pressure initiates a domino effect. Secondly, Halpern and the BIT intend to educate people about their biases. In both instances, this represents a hope that the market creates the change the BIT was founded to make happen. Needless to say, the market guarantees freedom of choice and decreases the likelihood of people having their inner econs disregarded.

It is in this context that nudge theory must be observed. It has undoubtedly and unnecessarily led to greater state involvement in the lives of people in Britain. It may be the case that TS were anticipating growth in intervention to change behaviour, and responding to it by developing a theory that they thought would appeal to both sides of the political spectrum. They were writing in response to popular rejection of behaviour change interventions, and designed an elegant, appealing

and novel solution that on the surface would appeal to both sides of the political spectrum.

In this nudge theory fails. Nudge theory can still be supported sincerely by the left, as there is enough evidence to suggest that it could be effective if libertarian concerns were disregarded, though the evidence is mixed for this as well, and it remains to be seen whether the lack of deliberative procedures hinders the BIT's durability. For conservatives, however, there is no reason to support nudge theory. The Conservative government saw no major reduction in state size (Daley, 2015), and the small one that did occur is difficult to credit to the BIT or nudge theory. With this experience, conservatives may in future consider the choice not between coercive behavioural policy and psychological behavioural policy as TS present, but between behavioural policy and non-interventionism.

8. <u>Conclusion</u>

The BIT was the first example of an institution claiming inspiration from nudge theory, but its lack of consistency with nudge theory indicates a few reasons to doubt its utility. The pattern is strong enough for it to not be explained by administrative reasons. The first finding is that the BIT often found non-psychological approaches more useful than psychological ones. The second is that the BIT found it difficult to remain within the ethical framework described in *Nudge*, and eventually abandoned any pretence of using this framework. Thirdly, there is a trade-off between effectiveness and freedom of choice when using psychology.

The BIT has fallen short of its own expectations in most trials, but exceeded the expectations of administrators and critics. This study has given reason to doubt the effectiveness of policy using behavioural science, at least when it remains within ethical boundaries, but the BIT remains optimistic. If the approach is to be successful it needs to fulfil Halpern's hope that small changes created by policy-makers will have a snowball effect. This will be an appropriate area for research, if enough data emerges to distinguish between government and market effects, perhaps using comparative and longitudinal studies. Australia might prove the most fertile ground for this, as some states have psychological policies and some do not. Opposite effects, perhaps arising from the erosion of deliberative reasoning abilities of proclivities as deliberative democrats prognosticate, would have to be considered as well. An implication of this study is that there is a shortage of nudges available. This could be tested to see if the experience is replicated elsewhere, and would be strongly vindicated if many of the policies in sister departments were duplicates.

The most important future study might be to test whether people are satisfied with the BIT's pallid ethical approach. Despite the reasons for conservatives to oppose psychological policy, and the relatively little use it has proven to be, it will continue to be used in both Britain and abroad. Having gained popularity based on its commitment to a libertarian ethos that now seems untenable, it has to be established whether Halpern's democratic approach could gain popular support. This could be compared to support for market nudges, or general awareness of the effect of biases on behaviour. Another potentially vital study is whether the organisations that the BIT is training to make policies, including frontline service providers, are using TS's, or any, ethical boundaries.

The move away from providing policy advice makes a mockery of claims about a psychological state. However, the psychological approach has gained wide, if thin, acceptance and usage in governance. This is an indication that conscious use of psychology in policy-making is likely to endure; public administration is more effective when administrators are aware of human psychological flaws. What the behavioural approach has provided is a few tools to make a few areas of government more effective. If this was the real aim, then the BIT, and nudge theory has been reasonably successful in this. By connecting government policy with a sensibility gained from behavioural science, the BIT has formally given government an inkling of a nuance that commerce has had for a long time, and increased the success of government in some areas where it generally fails.

Though it is wrong to condemn nudge theory as being without utility for policy-makers, there is every reason to be disappointed with the performance of both the BIT and nudge theory. The evidence does not suggest that nudge theory provides a third way that will unite the left and the right around a policy approach that will provide welfare gains while respecting freedom of choice. It managed to provide policy recommendations that do not rely on coercion, but still had little reservation about using its power to influence the thought and behaviour of subjects, even when the thoughts it sought to instil were a degraded version of the truth, and the behaviours not in the subjects' interest.

In abandoning the need to cater policies towards humans for the benefit of econs, the BIT has instead constructed policies to suit the leadership of the party that gained a plurality of votes at the previous election, and justified it in the same manner as political institutions normally do. Researchers could provide a service by asking whether nudge can in practice ever match the criteria in nudge theory, whether the experience

of the BIT in not living up to the moral restraints that legitimise nudging is repeated elsewhere and why.

Policy-makers would benefit by remembering that policies that seek behaviour change are not inevitable. Before the choice between coercive and non-coercive policies, there is a choice to intervene at all. Particularly, there are no compelling reasons for conservatives to support the BIT or nudge theory.. It is worth noting that in no instance has the BIT replaced coercive policies. The BIT was initiated by people who abhor large government and identified in nudge theory an opportunity to shrink the state while still pursuing the goals that governments pursued in the twentieth century. Nothing in the BIT's body of work suggests that it was working towards that aim, or that it did anything conducive to that. In every instance the BIT tried to make citizens conform to behaviours that the government wants them to have, restricted only by normal democratic concerns. There is therefore no suggestion that nudge theory can provide the basis of social policy for libertarians and conservatives. On the contrary, there is evidence that their traditional solution to policy problems, to rely on market forces, may provide better solutions to the same problems it was hoped the BIT would solve.

Reference List

Anderson, J. (2010). Nudge: Improving Decisions about Health, Wealth, and Happiness, Richard H. Thaler and Cass R. Sunstein. Yale University Press, 2008. x + 293 pages. [Paperback edition, Penguin, 2009, 320 pages.]. *Economics and Philosophy*, 26(03), 369–376. doi:10.1017/s0266267110000301

Barton, A., & Grüne-Yanoff, T. (2015). From Libertarian Paternalism to Nudging—and Beyond. *Rev.Phil.Psych.*, 6(3), 341–359. doi:10.1007/s13164-015-0268-x

Burgess, A. (2012). Nudging Healthy Lifestyles: The UK Experiments with the Behavioral Alternative to Regulation and the Market. *Eur. J. Risk Reg.*, *3*, 3.

Bradbury, A., McGimpsey, I., & Santori, D. (2013). Revising rationality: the use of "Nudge" approaches in neoliberal education policy. *Journal of Education Policy*, 28(2), 247–267. doi:10.1080/02680939.2012.719638

Brown, P. (2012). A Nudge in the Right Direction? Towards a Sociological Engagement with Libertarian Paternalism. *Social Policy and Society*, *11*(03), 305–317. doi:10.1017/s1474746412000061

British Broadcasting Corporation (2013, April 13). *'Mumbo-jumbo'*
personality tests for jobseekers criticised by Labour'. Retrieved
from http://www.bbc.co.uk/news/uk-politics-22356754

British Medical Association (2012). Behaviour change, public health and
the role of the state - BMA Position Statement. Retrieved online
from
https://www.bma.org.uk/collective-voice/policy-and-research/public
-and-population-health/behaviour-change

Coase, R. H. (1976). Adam Smith's view of man. *The Journal of Law &*
Economics, *19*(3), 529-546.

Corbett, S., & Walker, A. (2013). The big society: Rediscovery of 'the
social'or rhetorical fig-leaf for neo-liberalism?. *Critical Social Policy*,
33(3), 451-472.

Curtis, P. (2011, February 20). 'Nudge unit' not guaranteed to work, says
Oliver Letwin. *The Guardian*. Retrieved online from
http://www.theguardian.com/politics/2011/feb/20/nudge-unit-oliver-l
etwin

Daley, J. (2015, November 28). If Tories won't radically reduce the size
of the state, what's the point of them? *The Telegraph*. Retrieved
online from
http://www.telegraph.co.uk/news/politics/georgeosborne/12022841

/If-Tories-wont-radically-reduce-the-size-of-the-state-whats-the-poi
nt-of-them.html

Department for Environmental, Food and Rural Affairs (2008). *A
Framework for Pro-Environmental Behaviours* (*Ref*: PB13574 PDF).
Retrieved online from
https://www.gov.uk/government/uploads/system/uploads/attachme
nt_data/file/69277/pb13574-behaviours-report-080110.pdf

De Ridder, D. (2014). Nudging for beginners. *The European Health
Psychologist*, *16*(1), 2-6.

Di Nucci, E. (2013). Habits, Nudges, and Consent. *The American
Journal of Bioethics*, 13(6), 27–29.
doi:10.1080/15265161.2013.781711

Felsen, G., & Reiner, P. B. (2015). What can Neuroscience Contribute to
the Debate Over Nudging? *Rev.Phil.Psych.*, 6(3), 469–479.
doi:10.1007/s13164-015-0240-9

Gigerenzer, G. (2015). On the Supposed Evidence for Libertarian
Paternalism. *Rev.Phil.Psych.*, 6(3), 361–383.
doi:10.1007/s13164-015-0248-1

Grill, K. (2014). Expanding the Nudge: Designing Choice Contexts and
Choice Contents. *Rationality, Markets and Morals*, *5*, 139-162.

Grüne-Yanoff, T. (2012). Old wine in new casks: libertarian paternalism still violates liberal principles. *Soc Choice Welf*, 38(4), 635–645. doi:10.1007/s00355-011-0636-0

Halpern, D. (2015). *Inside The Nudge Unit*. London, UK. Penguin.

Hansen, P. G., & Jespersen, A. M. (2013). Nudge and the manipulation of choice: A framework for the responsible use of the nudge approach to behaviour change in public policy. *Eur. J. Risk Reg.*, 3.

Hausman, D. M., & Welch, B. (2010). Debate: To Nudge or Not to Nudge*. *Journal of Political Philosophy*, 18(1), 123–136. doi:10.1111/j.1467-9760.2009.00351.x

Hickman, M. (2011, January 3). Nudge, nudge, wink, wink... How the government wants to change the way we think. *The Independent*. Retrieved online from http://www.independent.co.uk/news/uk/politics/nudge-nudge-wink-wink-how-the-government-wants-to-change-the-way-we-think-217 4655.html

John, P. (2013). Policy entrepreneurship in UK central government: The behavioural insights team and the use of randomized controlled trials. *Public Policy and Administration*, 29(3), 257–267. doi:10.1177/0952076713509297

John, P., Sanders, M., & Wang, J. (2014). The Use of Descriptive Norms
in Public Administration: A Panacea for Improving Citizen
Behaviours? SSRN Journal. doi:10.2139/ssrn.2514536

Jones, R., Pykett, J., & Whitehead, M. (2013). The geographies of policy
translation: how nudge became the default policy option. *Environ.
Plann. C*, 32(1), 54–69. doi:10.1068/c1252

Junghans, A. F., Cheung, T. T., & De Ridder, D. D. (2015). Under
consumers' scrutiny - an investigation into consumers' attitudes
and concerns about nudging in the realm of health behavior. B*MC
Public Health*, 15(1). doi:10.1186/s12889-015-1691-8

Kosters, M., & Van der Heijden, J. (2015). From mechanism to virtue:
Evaluating Nudge theory. *Evaluation*, 21(3), 276–291.
doi:10.1177/1356389015590218

Leggett, W. (2014). The politics of behaviour change: nudge,
neoliberalism and the state. *Policy & Politics*, 42(1), 3–19.
doi:10.1332/030557312x655576

Lin, V (2003): Competing rationalities. In: Lin, V & Gibson, B:
Evidence-Based Health Policy. Melbourne: Oxford UP, 3-17

Loewenstein, G., Asch, D. A., Friedman, J. Y., Melichar, L. A., & Volpp,
K. G. (2012). Can behavioural economics make us healthier? *BMJ*,
344(may23 1), e3482–e3482. doi:10.1136/bmj.e3482

Marlow, M. L. (2014). *Weight loss nudges: market test or government guess.* Working Paper, Mercatus Center at George Mason University, Arlington, VA.

Martin, C. (2015). Nudging the public sphere: a Habermasian perspective on public deliberation as an aim of moral education. *Journal of Moral Education*, 44(4), 440–456. doi:10.1080/03057240.2015.1087392

McGonigal, K. (2012, June 1) The Bad Science of Banning Soda. *Psychology Today*. Retrieved from https://www.psychologytoday.com/blog/the-science-willpower/2012 06/the-bad-science-banning-soda

Mols, F., Haslam, S. A., Jetten, J., & Steffens, N. K. (2015). Why a nudge is not enough: A social identity critique of governance by stealth. *European Journal of Political Research*, 54(1), 81–98. doi:10.1111/1475-6765.12073

Moreira, M.T., Oskrochi, R. and Foxcroft, D.R. (2012) Personalised normative feedback for preventing alcohol misuse in university students: Solomon three-group randomised controlled trial. *PloS one* 7.9   e44120.

OLIVER, A. (2013). From Nudging to Budging: Using Behavioural Economics to Inform Public Sector Policy. *Journal of Social Policy*, 42(04), 685–700. doi:10.1017/s0047279413000299

Pagels, J. (2014, November 4) You're Just As Likely To Die En Route To Vote Than To Impact An Election Outcome, *Forbes*, Retrieved Online from http://www.forbes.com/sites/jimpagels/2014/11/04/youre-just-as-likely-to-die-en-route-to-vote-than-to-impact-an-election-outcome/#72ec57a47765

Parsons, W. (2002). From Muddling Through to Muddling Up - Evidence Based Policy Making and the Modernisation of British Government. *Public Policy and Administration*, 17(3), 43–60. doi:10.1177/095207670201700304

Science and Technology Select Committee (2011). Behaviour Change. 2nd Report of Session 2010-12. HL Paper 179. London. The Stationary Office.

Selinger, E., & Whyte, K. (2012). Nudging cannot solve complex policy problems. *European Journal of Risk Regulation*, 3, 26.

Sunstein, C. R. (2014). Nudging: A Very Short Guide. *Journal of Consumer Policy*, 37(4), 583–588. doi:10.1007/s10603-014-9273-1

Thaler, R. H. (2015) Foreword. In D. Halpern *Inside the Nudge Unit* (pp. IX - XIV). London, UK. Penguin.

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

Walsh, B. (2013, August 14). The Surprisingly Large Energy Footprint of the Digital Economy. *Time.* Retrieved online from http://science.time.com/2013/08/14/power-drain-the-digital-cloud-is -using-more-energy-than-you-think/

Whitehead, M., Jones, R., & Pykett, J. (2011). Governing irrationality, or a more than rational government? Reflections on the rescientisation of decision making in British public policy. *Environment and Planning A*, 43(12), 2819–2837. doi:10.1068/a43575

Yeung, K. (2011). Nudge as Fudge. *The Modern Law Review*, 75(1), 122–148. doi:10.1111/j.1468-2230.2012.00893.x

Young, W., & Middlemiss, L. (2012). A rethink of how policy and social science approach changing individuals' actions on greenhouse gas emissions. *Energy Policy*, *41*, 742-747.

Appendix

The policies contained in this list are all unique and distinctive policies contained in the BIT's output until July 8th 2016.

MINDSPACE

January 1st 2010

No policy recommendations. This was a re-hash of a report authored by the Prime Minister's Strategy Unit, and the date the BIT website says it was published predates the BIT.

Better choices: better deals, behavioural insights team paper

April 13th 2011

1. Implementation of 'Midata', online price comparison platform that collects individual consumer information.

Problem: Lack of price comparison for energy suppliers, caused by inertia, leading to lack of information.

Bias used: Overcomes and factors in inertia rather than uses it.

Inner Econ: The cloud mechanism's carbon footprint used for midata may make it a net loss for some users, and some will have privacy concerns, but it can be expected that most users match TS's criteria.

Choice: Does not restrict choice or alter incentives.

Transparency: Users will know what midata is for and how it works, but this is less important because a bias is not being used.

2. Annual credit card statements, issued by credit card companies including information on how to switch suppliers

Problem: Consumers lack information about financial survices and charges, presumably caused by their own inertia.

Bias used: None, because it changes the provision of information which is a rational market solution

Inner Econ: Yes for the nudgee, in this case the consumer.

Choice: Does not restrict choice or alter incentives for the consumer.

Transparency: Unimportant because a bias is not being used.

Behaviour change and energy use: behavioural insights team paper

July 6th 2011

3. Providing outlay for the Green Deal, and then using the money saved on future energy bills. Additional support given in the form of a holiday from council tax and retail vouchers.

Problem: Low uptake due to inertia and time discounting.

Bias used: None, it relies on rational incentives, but the report says it taps into loss aversion.

Inner Econ: Yes

Choice: Preserved, in the sense that it does not disincentivise any behaviours.

Transparency: Fufilled

4. Offer energy saving products at discounts varying according to how many people in a local area purchase them.

Problem: Low uptake, presumably due to inertia

Bias used: Susceptibility to social pressure/herd instinct

Inner Econ: Encouraging people to apply social pressure, and subjecting people to it from neighbours may violate this criteria.

Choice: Incentives can be altered and choice can be circumscribed by excess social pressure in this context.

Transparency: Easy for the nudgee to decipher.

5. Offering loft insulation as people move house, and use the RICS to recommend it. Additionally, offer of loft clearing in conjunction with loft insulation.

Problem: People do not take advantage of loft insulation policies due to inertia and inconvenience.

Bias used: By easing friction it bypasses people's laziness, authority bias.

Inner Econ: Yes

Choice: Preserved

Transparency: Fulfilled

6. Using the moving of home as a salient moment to communicate information about home energy usage, displayed on EPCs with the style of communication altered to emphasise some facts over others.

Problem: Lack of information about home energy usage.

Bias used: Anchoring, the tendency to better remember and to rely too heavily on one piece of information (often the first piece of information that is noticed) and attentional bias, but it is till largely reflective.

Inner Econ: Yes

Choice: It does not propel someone towards action, so no action is forbidden. However, because it is a change of mindset that is targetted by the apparent nudge, this effect is very difficult to avoid.

Transparency: In re-framing the matter transparency is not immediately evident.

7. Provide a report to households on domestic energy consumption, compared to both the neighbourhood average and efficient neighbours, demonstrated with a graph, data and a smiley face.

Problem: Inertia and lack of information about energy consumption.

Bias used: Re-framing of the issue, as well as social pressure and some debiasing effect.

Inner Econ: Information generally should accord with the inner econ if it is complete with relevant information. It relies somewhat on shame by comparing households with neighbours and depriving some of smiley faces.

Choice: The receipt of information is unavoidable, but no behaviour is forbidden or compelled.

Transparency: In re-framing the matter transparency is not immediately evident.

8. Change the design of EPCs to make it more simple and highlight smaller changes that can be made.

Problem: Inertia

Bias used: Removes friction

Inner Econ: Because the changes proffered benefit that the nudgee it does accord with the inner econ.

Choice: The receipt of information is unavoidable, but no behaviour is forbidden or compelled.

Transparency: Seems like the reader could figure this out.

9. Reduce the carbon footprint of government buildings by changing defaults in energy use, displaying information in buildings and holding competitions between buildings.

Problem: Inertia, lack of information.

Bias used: Relies on rational incentives mainly, but removes friction and provides new information anchors.

Inner Econ: This is a collective action, but it is done by people as employees who are paid to do it, so there is no issue.

Choice: Preserved

Transparency: Fulfilled

Fraud, error and debt: behavioural insights team paper

February 6th 2012

10. Altering tax debt collection letters to increase payments to include social norms about tax payment.

Problem: Underpayment of taxes, no reason is offered.

Bias used: Social norms relying on social pressure and an appeal to groupthink

Inner Econ: Unless the recipient is under threat of legal action, payment of taxes is a collective act, and using social norms an unwelcome pressure.

Choice: The receipt of communication is unavoidable, but no behaviour is forbidden or compelled.

Transparency: In re-framing the matter transparency is not immediately evident.

11. Letters sent to medics to increase tax compliance.

Problem: Underpayment of taxes, no reason is offered.

Bias used: Simplified letters and social norms depend on removing friction and social pressure

Inner Econ: Unless the recipient is under threat of legal action, payment of taxes is a collective act, and using social norms an unwelcome pressure.

Choice: The receipt of communication is unavoidable, but no behaviour is forbidden or compelled.

Transparency: In re-framing the matter transparency is not immediately evident.

12. Changing the letters sent to car tax evaders to include simplified letter and picture of the untaxed car on camera.

Problem: Underpayment of taxes, no reason is offered.

Bias used: Simplified letters and social norms depend on removing friction and social pressure

Inner Econ: Unless the recipient is under threat of legal action, payment of taxes is a collective act, and using social norms an unwelcome pressure.

Choice: The receipt of communication is unavoidable, but no behaviour is forbidden or compelled.

Transparency: In re-framing the matter transparency is not immediately evident.

13. Letters sent to resolve overpaid tax credits.

Problem: To collect debts, no reason is offered.

Bias used: The most successful letters used personalisation and framing the issue as a collective message.

Inner Econ: Unless the recipient is under threat of legal action, repayment of tax credits is a collective act.

Choice: The receipt of communication is unavoidable, but no behaviour is forbidden or compelled.

Transparency: In re-framing the matter transparency is not immediately evident.

14. Personalised text messages sent to collect fines imposed by courts.

Problem: No reason is offered.

Bias used: Egocentrism is the target of personalisation.

Inner Econ: Unless the recipient is under threat of legal action, repayment of tax credits is a collective act.

Choice: The receipt of communication is unavoidable, but no behaviour is forbidden or compelled.

Transparency: In re-framing the matter transparency is not immediately evident.

15. Countering fraud of the single person discount scheme by prompting honesty at the beginning of the form.

Problem: No reason is offered.

Bias used: Possibly anchoring and social pressure by holding people to their own norms.

Inner Econ: Unless the recipient is under threat of legal action, repayment of tax credits is a collective act.

Choice: Preserved

Transparency: PNudgee is probably unaware of the reason for the requirement.

16. Emphasising punishment in letters to plumbers who underpay tax.

Problem: No reason is offered

Bias used: None

Inner Econ: Using fear to generate tax payment is not good for anybody who recieves these letters.

Choice: The letter and threat are unavoidable

Transparency: Less important when no bias is being used.

17. Emphasising personal beliefs about taxation of the recipient of letters to companies who have underpaid tax.

Problem: People do not feel guilty enough about not paying tax.

Bias used: Seems to be perfectly reflective

Inner Econ: The effect is for the nudgee's companies, so it is not applicable, though presumably the individual is losing something in the process.

Choice: Preserved

Transparency: Recipients probably know that they are being told to think about their personal beliefs because the government wants their money.

Applying Behavioural Insights to Charitable Giving

May 24th 2013

18. Providing default donation options in communication urging monthly donation.

Problem: Thinking of an amount of money to donate involves too much friction for people with inertia, a very weak example of a bias.

Bias used: Inertia and anchoring on defaults.

Inner Econ: Certainly not, and research shows that people are uneasy being nudged when it will cost them money.

Choice: Present, but defaults are so effective that it must always be questioned whether they abuse inertia.

Transparency: Easy to understand

19. Changing the default option of automatic escalation of donations from opt-out to opt-in.

Problem: Insufficient money donated to charity, perhaps due to inertia, but a very weak example of a bias.

Bias used: Inertia

Inner Econ: Certainly not, and research shows that people are uneasy being nudged when it will cost them money.

Choice: The great increase in donations indicate that in this defaults did take advantage of inertia.

Transparency: Easy to consider that people did not understand the implication of the opt-in, or even notice it.

20. E-cards at Christmas from colleagues who already donate to charity, with a picture of the colleague, to encourage charitible giving.

Problem: Insufficient money donated to charity, perhaps due to inertia, but a very weak example of a bias.

Bias used: Susceptibility to social pressure.

Inner Econ: Certainly not, and research shows that people are uneasy being nudged when it will cost them money. Shaming tactics are not in accordance with this criteria either.

Choice: The communication, relying on shame, is unavoidable, and does alter incentives, perhaps to a large degree in some subjects.

Transparency: Some people will not be aware they are being shamed into making charitible donations.

21. Combining personalised messages with giving sweets to potential charity donors.

Problem: Insufficient money donated to charity, perhaps due to inertia, but a very weak example of a bias.

Bias used: Susceptibility to social pressure.

Inner Econ: Certainly not, and research shows that people are uneasy being nudged when it will cost them money.

Choice: Preserved, but the communication is unavoidable and it may engender shame in some instances.

Transparency: Some people will not be aware they are being shamed into making charitible donations.

22. Prompting people to leave money to charity in their wills, emphasising a social norm.

Problem: Insufficient money donated to charity, presumably due to inertia, but a very weak example of a bias.

Bias used: Susceptibility to social pressure.

Inner Econ: Certainly not, and research shows that people are uneasy being nudged when it will cost them money. Shaming tactics are not in accordance with this criteria either.

Choice: The communication, relying on shame, is unavoidable, and does alter incentives, perhaps to a large degree in some subjects.

Transparency: Some people will not be aware they are being shamed into making charitible donations.

Test, Learn, Adapt: Developing Public Policy with Randomised

Controlled Trials

June 14th 2013

No policy recommendations

Applying Behavioural Insights to Organ Donation

December 23rd 2013

23. Change the design of the landing page for the NHS's organ donation website

Problem: Insufficient organ donation rates, presumably due to inertia, but a very weak example of a bias.

Bias used: The reciprocity option was the most successful, and so social pressure is the bias used here.

Inner Econ: Organ donation is probably in accordance with this criteria. Social shame as a means of achieving on outcome is not always okay, but it does not seem like an imposition here.

Choice: The BIT just used the most effective option they have after trialling them, so they are taking a completely pragmatic approach to overcoming faculties of reason and overriding choice.

Transparency: It will be clear that the website is trying to encourage organ donation, but not that the process was designed so ruthlessly to achieve it.

Growth vouchers

February 4th 2014

No policy recommendations

EAST: Four Simple Ways to Apply Behavioural Insights

April 11th 2014

No policy recommendations

Clinical Judgement and Decision-Making in Children's Social Work: An analysis of the 'front door' system

April 15th 2014

24. Develop quantitative predictive modelling to identify effective practices.

Problem: Social workers have a poor grasp of evidence, so presumably many biases can be present among them.

Bias used: A heuristic is provided, overcoming biases by replacing it with a more effective one.

Inner Econ: If successful.

Choice: Unimportant as the people in question are employees

Transparency: Fulfilled

25. Introduce feedback loops to help social workers learn from past decisions.

Problem: Lack of information and a resistance to a growth culture.

Bias used: Bias is overcome rather than used.

Inner Econ: If successful

Choice: Unimportant as the people in question are employees

Transparency: Fulfilled

26. Improve the inputs to the system, by developing simpler systems for filtering out information.

Problem: Too much irrelevant information relayed between social workers as colleagues and with the police, so the sub-optimal outcome is the result of limited cognitive ability, and could be caused by many biases.

Bias used: A heuristic is provided, overcoming biases by replacing it with a more effective one..

Inner Econ: If successful

Choice: Unimportant as the people in question are employees

Transparency: Fulfilled

27. Development of heuristic tools and/or checklists to guide decision-making without the complexity of actuarial tools.

Problem: Room for human error and misjudgment due to lack of awareness of data and complexity.

Bias used: Heuristic provided.

Inner Econ: If successful

Choice: Unimportant as the people in question are employees

Transparency: Fulfilled

Reducing Mobile Phone Theft and Improving Security

September 1st 2014

No policy recommendations

The Behavioural Insights Team Update 2013-2015

July 23rd 2015

28. Used 'implementation intentions' to get specific commitments from jobseekers to achieve set goals.

Problem: Inertia leads to inactivty from unemployed jobseekers

Bias used: perhaps anchoring, but it seems to depend largely on appealing to rationality.

Inner Econ: Yes, aside from people who prefer unemployment, and those who cannot get jobs for whom extra effort is a waste of energy.

Choice: As the jobseeker's sole source of income is jeopardised, this alters incentives significantly.

Transparency: Jobseekers will be aware that this is a psychological action.

29. Using text messages to generate attendance at recruitment fairs, drawing on reciprocity.

Problem: Inertia leads to less attendance at recruitment fairs than the authorities hoped.

Bias used: Reciprocity was the mechanism.

Inner Econ: Mixed because of the use of shame and people who prefer unemployment, and those who cannot get jobs for whom extra effort is a waste of energy.

Choice: Preserved, but by engendering shame incentives are altered somewhat.

Transparency: Jobseekers may not be aware that this is a psychological action.

30. Simplifying and personalising government emails to small businesses about beneficial programs available to them.

Problem: Insufficient uptake of those programs, presumably as a result of inertia.

Bias used: As it assumes inertia and seeks to make the process simpler the bias used also seems to be inertia, while personalisation depends on egocentrism.

Inner Econ: Yes.

Choice: Preserved.

Transparency: Subtle but not indiscernable.

31. Simplification of prescription forms to reduce error rates.

Problem: General cognitive deficiencies result in lots of problems caused by biases.

Bias used: Making the forms clearer and easier to use does not use a bias, it helps to overcome inertia.

Inner Econ: Yes

Choice: No choice is left, but it is completely unimportant because the nudge is directed at employees.

Transparency: Fulfilled.

32. Placing anti-smoking stickers on pregnancy tests.

Problem: Addiction is the result of several biases that smoking can be attributed to, especially attentional bias.

Bias used: Attentional bias by placing the stickers at a salient place and moment.

Inner Econ: Social shame as a means of achieving on outcome is not always okay, but it does not seem like an imposition here.

Choice: Preserved

Transparency: Fulfilled

33. Counter resistance to antibiotics by reducing their unnecessary prescription, to be achieved through posters and leaflets informing doctors of norms.

Problem: Discounting the future effects of prescriptions in the moment.

Bias used: An example of persuasion rather than using a bias.

Inner Econ: Yes

Choice: Preserved

Transparency: Fulfilled

34. Enhance efforts to promote smoking cessation annually in October by removing clutter from the 'Stoptober' website.

Problem: A website visually unappealing to the biases, changed by insights gained from nudge 23 and inertia in smoking cessation efforts.

Bias used: A preference for a simple and attractive website, not really a bias used though can be conceived of as removing friction.

Inner Econ: Yes

Choice: Preserved

Transparency: Fuflilled

35. Prompting the switching of energy suppliers on the envelopes of letters of the winter fuel allowment.

Problem: Inertia

Bias used: Attentional bias by placing the stickers in a salient position.

Inner Econ: Yes

Choice: Preserved

Transparency: Fuflilled

36. Provide performance incentives for the pupil premium, designed to benefit school students from disadvantaged backgrounds.

Problem: No bias present

Bias used: No bias used it appeals to rationality

Inner Econ: Yes

Choice: Unimportant because the target of the recommendation is a government employee acting in this position.

Transparency: Fuflilled

37. Adopting a friendlier tone to ethnic minority applicants at a salient moment during applications to the police force, as well as stressing the importance of their application to their ethnic community.

Problem: Inertia

Bias used: Attentional bias by making the appeal at a salient moment, but it seems to appeal to rationality.

Inner Econ: Yes

Choice: Preserved

Transparency: Fuflilled

38. Encourage active decision-making about the immigration status of prospective employees for businesses.

Problem: Inertia and time discounting.

Bias used: This really just plays upon the fear of managers.

Inner Econ: Almost certainly not, it directly attacks a freely made voluntary exchange, and there is very little chance that the manager or company is being saved from a greater punishment.

Choice: Preserved

Transparency: Fufilled

39. Increasing the number of army reservists by emailing people with a declared interest and addressing it from a real officer.

Problem: Inertia if the prospective soldiers do not fulfill an earlier intention.

Bias used: A kind of anthropomorphism.

Inner Econ: There could be a reverse time discounting present if the long term risks of being a soldier outweigh the immediate benefits.

Choice: Preserved

Transparency: Some recipients may miss the fact that they are being manipulated.

40. Worked towards personal cyber security by encouraging people to create passwords from three strung together words.

Problem: Inertia and lack of information about cyber security.

Bias used: Provides advice and a very narrow choice architecture, so it removes friction.

Inner Econ: Yes

Choice: Preserved

Transparency: Fuflilled

41. Displaying the fiscal costs on energy labels of washers, dryers and washer-dryers to encourage the buying of energy efficient goods.

Problem: Inertia, lack of information and time discounting

Bias used: Attentional bias by placing the stickers in a salient position, but is largely an appeal to rationality.

Inner Econ: Yes

Choice: Preserved

Transparency: Fuflilled

42. Using boiler engineers to promote the saving of energy to homeowners in person and by leaving notes.

Problem: Inertia leads to energy inefficient behaviours.

Bias used: Biases towards authority and 'trusted messengers'.

Inner Econ: Yes

Choice: Preserved

Transparency: Some people particularly deferential to authority may not see the nudge, but because it was ineffective there does not seem to be a strong risk of this.

Behavioural Insights and the Somerset Challenge

July 31st 2015

43. Promote the lifestyle aspects, rather than the long term benefits, of university to encourage students to apply.

Problem: Inertia, time discounting

Bias used: Time discounting

Inner Econ: Yes

Choice: Preserved

Transparency: Fuflilled

A head for hiring: the behavioural science of recruitment and selection

December 8th 2015

No policy recommnedations

Evaluating Youth Social Action - Final Report

January 14th 2016

No policy recommendations

Making the change: Behavioural factors in person- and community-centred approaches for health and wellbeing

March 22nd 2016

No policy recommendations

Reducing Mobile Phone Theft and Improving Security - Paper 2

March 23rd 2016

No policy recommendations

Smart Meters Derogation Guidance: supporting energy supplier applications for trials of in-home display alternatives

April 27th 2016

No policy recommendations

Applying behavioural insights to regulated markets

May 26th 2016

No policy recommendations

Decision-making in children's social care: quantitative analysis

July 8th 2016

No policy recommendations