carr
centre for analysis
of risk and regulation

centre for analysis
of risk and regulation

Rational tools of government in a world of bounded rationality

Martin Lodge and Kai Wegrich

DISCUSSION PAPER NO: **75**
DATE: **December 2014**

LSE THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Rational tools of government in a world of bounded rationality

Martin Lodge and Kai Wegrich

Contents

Published by the Centre for Analysis of Risk and Regulation at the
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
UK

ISBN 2049-2718

Rational tools of government in a world of bounded rationality

Martin Lodge and Kai Wegrich

Abstract

Nudge, and the wider behavioural economics agenda, have become one of the most dominant themes in contemporary public policy. This paper points to a paradox underlying Nudge, namely, whether an approach that offers a rational approach towards bounded rationality can escape the constraints of bounded rationality itself. Firstly, this paper argues that governments and individual decision making are characterised by bounded rationality. Secondly, it considers how ways of organising Nudge inside executive government introduce their own biases; and, thirdly, how Nudge encourages particular unintended consequences. In conclusion, this paper suggests that it is time for Nudge enthusiasts to consider their own decision making biases and 'limits to rationality'.

Introduction

A spectre has been haunting practitioners and students of executive politics and governance – the spectre of Nudge. Far beyond Europe, the US and Australia, the Nudge agenda has been enthusiastically endorsed by governments of all colours and international organisations as a policy tool of choice. As any social movement, Nudge can also rely on its own manifesto, namely *Nudge* by Richard Thaler and Cass Sunstein (2008). Nudge may officially represent just one of a whole plethora of instruments seeking to shape the behaviour of individual, but it has also come to stand for the much wider fascination in behavioural economics in government (see Oliver 2013).

Nudge embraces an understanding of human decision making that is defined by bounded rationality i.e. the notion that decision making is characterised by constraints on individuals' cognitive capacities and resources in processing, often limited, information. It therefore represents a natural bed-fellow of the 'behavioural revolution' that has gripped the social sciences over the past few years (see also Amir and Lobel 2008). Nudge resonates with the contemporary fascination in risk and loss aversion (Kahneman and Tversky 1979), decision making biases (Ariely 2008, Gigerenzer 2014; Kahneman 2011) and the politics of 'happiness' and cognitive behavioural therapy. Policy researchers and economists have proclaimed behavioural economics as a new discovery regardless of the long-standing interest in bounded rationality in the social sciences (see Simon 1997).

Furthermore, the *Nudge* book is of particular interest given the politically and academically exposed position of its authors, especially Sunstein (as Obama's regulation supremo until 2012). In the UK, the part-privatised Behavioural Insights Team (or Nudge Unit) has played a similarly influential role at the centre of the UK government and on the international conference circuit. In addition, Nudge offers the opportunity to bring together those who believe in non-interventionist governance and those seeking to realise 'optimal' outcomes through intelligent design. It therefore combines the promise of 'cheap government' in an age of depleted financial resources with the promotion of 'choice' in an age of increasingly heterogeneous societies that no longer tolerate 'one size fits all' policies. Finally, Nudge also offers the illusion of cheap government in that once the right nudge has been chosen, individuals will choose optimal solutions by themselves without requiring costly enforcement activities.

This paper challenges the Nudge agenda by pointing to a basic paradox. It asks whether an approach that offers 'rational' policy making to address problems resulting from bounded rationality can overcome the limits of bounded rationality itself. In doing so, this paper deviates from the range of critical responses to the Nudge movement. One strain of criticism has focused on its philosophical foundation in 'liberal paternalism' (Rizzo & Whitman 2008). A further strain has argued that nudging is an act of non-transparent manipulation where accountability structures are deliberately left unclear (Rizzo & Whitman 2009; Wilkinson 2013). Nudging is therefore seen as inherently problematic when assessed in the light of basic principles of liberal democracy. A third strain, applying to behavioural

economics' informed approaches in general, is to attack a perceived 'theory-free' reliance on a view that 'the data speaks for itself' (Harford 2014).

This paper develops a different line of argument by dealing with the rationality assumptions of Nudge. Nudging assumes a world in which individuals are boundedly rational: decisions are affected by confirmation biases, loss aversion, and openness to optical manipulation. People make sub-optimal choices due to the perception of too high transaction costs; it is therefore assumed that individuals would inherently prefer to improve their choices. In other words, sub-optimal choices are a result of 'reasoning failure'. For example, individuals fail to register as organ donors, even when they are generally in favour of doing so (Abadie & Gay 2006). Similarly, they fail to sign up to health-care or pension plans, despite the realisation that basic coverage is likely to be insufficient in old age (Choi el al. 2001; Madrian & Shea 2000). They consume sugar- and calorie-rich drinks despite knowing about their content. Or, in the context of lesser developed countries, they are reluctant to allow their children to enjoy regular schooling due to short-term economic need, despite knowing the importance of education for their children. In sum, short-run benefits drive out much higher long term benefits, or distrust and lack of information make seemingly irrational behaviour rational. In addition, Nudge has been used to support compliance and cheapen enforcement; it is argued that by personalising messages and by exploiting social norms, individuals are more willing to pay taxes and parking fines on time or with less delay.

Such sub-optimal choices, or 'reasoning failures', are a product of underlying choice architectures. Nudge devices change the basic settings of the so-called choice architecture. The latter might be the result of previous policy decisions, or they may be the consequence of market incentives. Any social situation contains choice architectures and therefore nudges. By changing the choice architecture, humans are put into a position to act according to their preferences without facing major opportunity costs. They continue to satisfice, but achieve more desirable outcomes, for themselves and for wider society. In addition, they are usually not required to choose how to obtain the desired outcome; in the world of Nudge, individuals are free to 'opt out' rather than 'opt in'. Thus, individuals eat more healthily when encountering appealing fruit and vegetables rather than chocolate bars or bags of crisps, they cause less splatter when incentivised to undertake target practice in male urinals, and they happily donate organs. They send children to school as attendance is linked to cross-linked support packages.

While fiddling around with choice architectures may be regarded as being highly paternalistic (by those believing in the decision making competence of individuals) or as too individualistic (as it believes in the power of nudge without considering the power of business to manipulate and pressurise), it has much wider implications. It suggests that those deciding on choice architectures and on dominant psychological mechanisms are equipped with perfect rationality (Rizzo & Whitman 2008). Nudge assumes that the benevolent rational decision maker can identify 'rational' behaviour and therefore also sub-optimal choices arising from bounded rationality. Decisions are 'evidence-based', often backed up by 'gold standard' random control trials (John et al. 2011; John 2011).

Interventions are justified through 'paternalism'; nudged choices are claimed to reflect the 'real' preferences of nudged individuals. Thus, decision makers have the evidence base to make decisions as to how individuals can be manipulated. They are equipped with the persuasive authority to convince others regarding the superiority of nudging relative to other interventions. They are able to address potential interaction effects with other policies that apply to a particular problem, and they can make sense of the multiple motivations that apply to human behaviours. In other words, at the heart of nudge is a basic paradox. It assumes bounded rationality, but offers a 'comprehensive' vision of rationality to address problems caused by bounded rationality.

This paper explores this 'rationality paradox' at the heart of Nudge in three steps. First, we suggest that governments' decision making is boundedly rational. Such limitations affect policy interventions in general and are therefore not specific to Nudge. However, they are specifically problematic for Nudge given Nudge's argument that it is aware of bounded rationality. Then, we explore in more detail how bounded rationality affects the rationality of Nudging. We do so by looking at how the introduction of Nudge is faced with the limits of bounded rationality in a setting that is characterised by organisational and political logics on the one hand, and how Nudge is confronted by limits of bounded rationality in individual decision making on the other. In doing so, we are not developing an empirical argument given the limited and problematic evidence base. Instead, by highlighting the basic paradox at the heart of Nudge, we are primarily concerned with the rationale of Nudge.

The rationality 'paradox'

Nudge assumes that individuals and organisations have the capacity and motivation to change. It assumes that we know what people want. The capacity demands are small in the eyes of nudge enthusiasts as individuals simply follow paths chosen for them. In this sense, Nudge is different from those tools that seek to bring rationality into decision making, whether this is via performance management systems, cost-benefit analysis or rational budget programming systems. Nudge does not seek to reduce irrationality in government decision making through procedures; in contrast, it seeks to reduce irrationality by exploiting irrationality at the level of the target of public policy, the individual.

The attraction of Nudge is based on being both familiar and seemingly novel at the same time. It is familiar in that it recognises the bounded rationality-induced limitations in human behaviour that the enforcement and implementation literatures have been emphasising for some time. It is novel in that it places its emphasis on information and incentives as a tool to manipulate individual rather than organisational behaviours. As such, it fits a social science agenda that seeks to discover human psychology or to promote Cognitive Behavioural Therapy at the expense of others. It also fits a political agenda that pretends to be evidence based in order to find seemingly low-cost high intelligence measures.

Before exploring the consequences of the rationality paradox at the heart of 'nudge' and other behavioural economics-informed initiatives in executive government, it is critical to consider the justification for proposing the rationality paradox in the first place. After all, governments may not be suffering from bounded rationality in general, or, somewhat differently, may suffer from other dysfunctional symptoms that have nothing to do with bounded rationality.

The first line of argument to claim that governments are more rational than individuals relies on the superior resource base of governments. Governments possess processing capacity that go far beyond those of individuals. These superior resources involve:

- finance (ordinary individuals do not have the same amount of cash to spend on research and espionage);
- knowledge (individuals do not have the capacity to access networks of expertise to offer insight as to what to do with certain threats);
- organisation (individuals do not have armies of sitting and standing bureaucrats, consultants and academics at their disposal to gather information, to analyse it, and effect choices); or
- authority (individuals usually have not got the legitimacy to prohibit or permit certain activities).

In addition, governments are, within limits, able to constrain their behaviour through the generation of procedural safeguards (such as impact assessments) and the establishment of particular organisations (e.g. central banks to reduce the time inconsistency problem in monetary policy). In other words, the key characteristics of bounded rationality, the resource limitations on individuals that lead to potentially sub-optimal choices, are overcome through superior resources and procedural and other safeguards; governments do not suffer 'reasoning failures' as individuals do.

The second line of argument to argue that governments do not suffer the kind of bounded rationality that individuals do suggests that governments' limitations in their decision making cannot be treated in the same way as the limits on individual decision making. Governments are less prone to information asymmetry and processing problems. Failure, i.e. sub-optimal choices or 'blunders', emerge due to the exercise of 'power'. For example, because of the election-seeking behaviour of politicians, the successful capture of policies by concentrated interests, or other kind of distortions that might occur in the policy process. Policy making, according to this perspective, is about strategic and rational actors. Their interactions may lead to flawed outcomes, but these are due to institutional incentives or interest group constellations, not about bounded rationality per se. Somewhat relatedly, the study of 'blundering governments', such as the one presented by Anthony King and Ivor Crewe (2013), is largely about inherent weaknesses of the British policy making machinery rather than inherent limits to knowledge and administration.

Neither of these two lines of argument offers a real challenge to the view that governments are similarly bounded in their rationality as are individuals. Individual actors and actors inside government, and those trying to access government, pursue strategic aims. These actors, however well resourced, are boundedly rational (Miller 2000; Pierson 2000; 6 2014). Actors are myopic – resource-dependent on other actors; seek to protect their organisational and individual turf and reputation otherwise known as 'intended rationality' (Jones 2003: 397); display a tendency towards risk aversion and negativity bias; deal with multiple audiences both within and outside their own organisation; and are attached to particular values, symbols and rituals (March & Olsen 1983; Stark 2014). Both deal with uncertainty and the need to rely on heuristics which need not be inferior under conditions of uncertainty (see Gigerenzer 2014). Individuals as governments interact with others whose capacities and motivations are not always easy to observe or estimate. Indeed, the notion of 'reasoning failure' suggests that Nudge assumes well intentioned, but poorly informed individuals. However, much government activity is about dealing with the ill intentioned, and often, well resourced. Both individuals and governments face so-called wicked issues on a day-to-day basis, therefore further highlighting inherent cognitive and value-based trade-offs that characterise decision making.

Such a set of claims is hardly novel in the study of public and private organisations. For example, Simon (1997) notes how organisations face considerable information processing and decision making challenges. The metaphor of 'garbage can decision making' has also signalled the limits of 'rational' policy making: actors pay partial attention, attendance is fluid, preferences are unclear, and 'solutions are searching for problems' (Cohen et al. 1972). Informed by the view that governments' activities are best characterised as being shaped by bounded rationality, others have highlighted the partial attention that governments pay to any one issue at particular times (Baumgartner & Jones 1993; Jones 2003), how default policy settings have their own 'policy inheritance' effect (Rose 1990), and how the 'logic of appropriateness' shapes individual and organisational behaviours and decisions (March & Olsen 1983). Furthermore, organisations' learning is said to be biased towards confirming rather than challenging existing ways of doing things (March et al. 1991). Christopher Hood (1976) has also highlighted limits of administration that are part and parcel of the boundedly rational nature of government; some of the limits are due to inevitable time lags, others about genuine uncertainty, and others about costs of acquiring and searching for information. The standard litany of problems in executive government, such as control, co-ordination, or implementation are about boundedly rational actors pursuing their strategic objectives in the context of limited information or genuine uncertainty. Issues of 'multi-organisational sub-optimisation' (Hood 1976) emerge as organisations dealing with standard setting, information gathering and behaviour modification operate in dispersed arenas, are concerned about their reputation and focus on those activities that are of immediate importance to them.

Furthermore, even though governments have arguably more resources on tap than most individuals or networks of individuals, this does not mean that reasoning failure cannot occur in government. For example, the search for evidence-based

policy making is inherently about limited searches; trials, however gold standard, incorporate value and methodological choices in addition to political priorities. Whatever effort may have gone into designing gold standard-type experiments, the actual results from experiments are, at best, trivial. Experiments themselves are often based on flimsy foundations as social experiments are more complex than the (supposedly) controlled environment of medical trials – even if one tries to ignore the highly problematic selection biases that underpin medical trials (Kingston 2012): can undergraduates who are paid to spend some time in 'laboratories' really be seen as real-life equivalents? Can real-life trials overcome challenges in terms of scale, time, contamination effects and such like?

For both individuals and governments, intentional action can therefore quickly lead to unintended consequences. Unintended consequences (that are not necessarily undesirable) emerge from a range of sources (Merton 1936: 900). Bounded rationality, for Merton, is about high transaction costs,[1] dealing with limited knowledge, error prone assumptions and an 'imperious immediacy of interest'. The latter is defined by short-term interests driving out long-term concerns, considerations about interdependencies, moral implications and the possibility of interaction effects such as counter-learning (see also Hood 1976; Sieber 1981). Similarly, Charles Lindblom (1959) notes how resource constraints made any strategy that was not incremental both normatively and practically undesirable. Again, as the authors of these works suggest, these decision making biases apply to both organisational and individual decision making.

This is not to say that trialling policy interventions is meaningless, or that all government action will inevitably lead to unfavourable results. However, it suggests that actors in government are similarly boundedly rational as the individuals whose decision making they seek to manipulate. In other words, individuals and governments suffer from 'reasoning failure' (as termed by Nudgers). Individual as well as government decision making is inherently about transaction costs, it is about uncertainty, it is about error prone assumptions, and about short-term biases, and intended rationality. Both governments and individuals make decisions under conditions of bounded rationality, and therefore the consequences of the rationality paradox at the heart of Nudge deserves greater attention.

Organisations, tool choice and bounded rationality

In an ideal setting, government agencies would rely on 'evidence-based' Nudge-informed strategies, and they would avoid policy 'knee jerks' to media-feeding frenzies (Hood & Lodge 2005), and carefully consider costs and benefits of various regulatory options. Bounded rationality in an organisational and inter-organisational context stands in the way of adapting supposedly superior policy strategies given path dependencies, established constituencies, jurisdictional turf

[1] Merton (1936: 900) named these 'the *economic* problem of distributing our fundamental resources'.

battles between organisations, and confirmation bias that puts a premium on default strategies. This political context is hardly specific to Nudge. However, supporters of Nudge, and evidence-based policy making more generally, pay little attention to these sources of 'irrationality', or how these generic factors in executive government may impact on the Nudge agenda itself.

How, then, is Nudge organised within the context of executive decision making that is characterised by both inertia and knee-jerking? And, more importantly, is it able to withstand the kind of organisational processes that are associated with bounded rationality? One typical strategy for any reform approach has been the creation of special units at the heart of government that are supposed to advance a particular agenda. As noted, the UK Conservative-Liberal Democrat government established a 'nudge unit' inside the Cabinet Office, officially named the Behavioural Insights Team, set up in June 2010 after the general election and change in government. This unit was supposed to advocate Nudge thinking across government, whether it was in the way in which regulatory reform proposals might be recast, charity payments could be increased at a time when public money for charities was being slashed, healthy school meals were encouraged, or in the way in which letters to tax offenders were written to reduce delay in payments (see also Behavioural Insights Team 2014).[2] The second orthodox approach has been to force all decision making through procedural methodologies and thereby force nudging onto the agenda for policy making. Such hardwiring through deck-stacking forces decision makers to confront particular options.

These two orthodox ways of trying to integrate high intelligence policy ideas into the daily low intelligence life of executive decision making have usually led to limited results, because of Merton's 'imperious immediacy of interest'. This 'imperious immediacy of interest' can be separated into four key mechanisms: loose coupling, marginalisation, incrementalism and decomplexification. These mechanisms are not mutually exclusive, nor are they jointly exhaustive, but they capture a substantial variety of potential dynamics that lead to unintended effects. Nor are these mechanisms exclusive to Nudge, but apply to policy interventions more generally. Table 1 summarises these mechanisms.

Table 1 Organising Nudge in executive government

Loose coupling Lack of penetration into existing policy	*Incrementalism* Small scale change to existing policies
Marginalisation Attention crowded out by more urgent matters	*Decomplexification* Over-simplification leads to 'pure' solutions that fail to reflect complexity

[2] The Behavioural Insight team 'corporatised' itself in 2014, joining the charity Nesta and being part-owned by the charity itself, the staff and the government.

Loose coupling refers to a lack of co-ordination between different organisational logics that exist within government. Different units and departments within government have varied views as to their priorities and limited resources. They also develop distinct sectoral identities given their frequent exchanges with key constituencies. As a result, the politics within executive government are defined by the struggle between dispersed units, anxious to maintain their autonomy. It is therefore unsurprising that any attempt at imposing an organisational solution onto such dispersed setting will be received with scepticism, if not rejection. For example, Nudging will hardly appeal to civil servants in energy portfolios who are in close relational distance with large-scale energy firms. Having to 'nudge proof' policy initiatives will be seen as unwelcome to anyone, whether they are working in areas such as shop opening times, maternity leave or bovine tuberculosis. Similarly, forcing procedural devices into decision making processes is unlikely to automatically raise the profile of Nudge in government. Compromises between and within departments are more likely to be about budgetary allocations and the carving up of jurisdictions to address the bare necessities of ministerial reputation and blame management. Such a setting is not necessarily open to the introduction of procedural devices to force Nudge on to the table. This is particularly the case when Nudge czars in government are anointed from the ranks of junior civil servants whose eyes are on the quick ascent up the career ladder, not on picking fights with more senior officials in their own or a different department. Outside appointments to such units usually require the backing of very senior politicians to be granted any audience within government. In other words, the normal organisational life within government allows only for a loose coupling of the Nudge agenda to the 'real world' of decision making unless some political heavyweight forces the agenda on to reluctant parts within the executive. However, once that heavyweight has found a different playground, has been promoted upwards or sideways, or has bitten the proverbial dust, Nudge is likely to bounce back into its loosely attached status.

Marginalisation defines a process that leads to a similar outcome to loose coupling, but the underlying mechanism is different. Here, the proposed intervention is just one of many other important issues that decision makers have to consider. Therefore, Nudge gets marginalised as other priorities take over, whether this is because of the value basis of much policy making, the ambiguous evidence that allows for a range of evidence-based strategies to be considered, or that politicians prefer 'visible' policy strategies to achieve credit-claiming media headlines. A strategy that relies on non-transparent manipulation of peoples' preferences, whether it is stickers given to schoolchildren to reward them for 'healthy' meal choices, or the insertion of carefully worded sentences into official government communications, is hardly the kind of material that allows politicians to blow their own trumpet. Nudge is less likely to suffer from marginalisation in those areas where the stakes are particularly low – when both the costs and benefits of regulation are widely distributed across constituencies. But those are arguably areas where intervention might be least important in the first place.

Incrementalism in decision making in government also stands in the way of a comprehensive introduction of Nudge into policy making. After all, Nudge is

supposed to be based on evidence-based decision making, preferably generated through the gold standard random control trial. Incrementalism takes the default setting as given and centres on decision making at the margin. The introduction of Nudge as a serious endeavour calls for a questioning of the 'default option', which is likely to attract considerable resistance and opposition, thereby reducing the scope to achieve an agreement in the first place. For example, explicit attempts at wiping the existing stock of policies clean, such as the UK Red Tape Challenge initiative where all departments were to cut their stock of legislation and regulation unless a good reason could be found to retain them, eventually turn incremental as any comprehensive root or stock review over-stretches the capacities of government units (Lodge & Wegrich 2014). In particular, the demand to separate values (political objectives) from means (policy tools) increases the analytical complexity of such an exercise. Moreover, in a multi-actor setting, decision making is often easier to achieve when incremental steps are taken on the basis of an agreement on the means, but without requiring consensus on the underlying objectives or ends (Lindblom 1959). Nudge represents an approach that requires an agreement on both the ends and the means.

Finally, nudging is also problematic as it leads to *decomplexification*. In many ways, all bureaucracy is about decomplexification in the sense of creating categories and classifications to allow for a processing and 'normalisation' of the daily noise that government departments are exposed to. However, here *decomplexification* goes further, namely in the sense that nudging reduces the capacity of problem solving per se. As argued by Baldwin (2010), to address complexity one needs to encourage 'clumsy' and hybrid solutions and not search for 'elegant' ones. Similarly, the better regulation literature notes that regulatory problem solving should encourage combinations involving 'soft' regulation, self-regulation with incentives and, occasionally, command-and-control regulation. The enforcement literature, too, encourages a mixing of soft and hard instruments, and emphasises that it is only the presence of hard instruments that allows the functioning of soft instruments. Hybridisation and clumsiness are problematic to design, but they highlight that Nudge is, in principle, an 'elegant' solution that stifles creative combinations through its emphasis on particular forms of 'evidence' and on the calculation of costs and benefits. This provides policy bureaucrats with incentives to come up with regulatory designs that are simple and pure in their approach, rather than complex combinations of different tools and approaches.

In short, the 'imperious immediacy of interest' in organisational decision making provides a problematic setting for Nudge to succeed as a priority or to be informed by a higher degree of 'rationality' than other initiatives. Confirmation bias, risk aversion, turf battles and disproportionate information processing characterise politics and decision making at the top – hardly the kind of conditions that make evidence-based nudge units a solution to reasoning failures in executive government. In short, the bounded rationality conditions that define organisational decision making have not been transformed by the rise and rise of the Nudge agenda. Nudge's popularity is a product of bounded rationality in executive government. It is shaped by bounded rationality in decision making processes, and it is its itself developed by individuals who are boundedly rational. Such conditions

have a strong impact on the way in which Nudge is being developed and utilised, and the way its intended and unintended consequences emerge. As noted, the bounded rationality affecting government decision making is nothing that is specific to Nudge; what is specific about Nudge is that this is an approach that emphasises bounded rationality, but does little to acknowledge these limitation in its own approach.

Individuals, tool choice and bounded rationality

The above discussion is unlikely to raise opposition among enthusiasts of Nudge. They would suggest that their prescriptions were never meant to be universal, and that Nudge does offer some political benefits, namely, a reliance on low cost political initiatives that may free up space for engagement in more electorally appealing areas. Nudge-based interventions have had some effects, according to advocates' websites at least. At the same time, Nudge has had no noticeable effect in other areas, such as in food labelling (OECD 2010: chapter 6). This section does not seek to weigh the evidence supporting or disputing Nudge as a policy tool. Nor does it consider what kind of value judgements should underpin Nudge-type decisions. For example, asking individuals to 'opt in' so that they can access pornography on their laptops in their bedrooms is clearly a political choice as to whether and where Nudge should be employed. Similarly, the decision to rely on Nudge to deter certain behaviours or consumption patterns rather than to fall back on punitive tax levels is a political choice.

This section considers whether bounded rationality can be 'rationally' manipulated or whether the type of limitations noted by Merton over 75 years ago are applicable to Nudge as well. Such a question may sound puzzling as Nudge is exactly about exploiting those limitations. However, this requires a degree of superior knowledge about people's choices that may not always be present. As in the previous section, we note four ways in which bounded rationality may trump the best policy intentions: classification error, aggravation, placation and over-commitment. These are summarised in Table 2 (see also Sieber 1981, who adds functional disruption, goal displacement and exploitation to the list of reverse effects).

Table 2 Overview of unintended consequences

Classification error Wrong choices about which issues to nudge or regulate	*Placation* Nudges lead to intended bhaviour change, but fail to address and hide build-up of major problem
Aggravation Responses encourage opposite behaviour	*Over-commitment* Disappointment effect after over-enthusiastic uptake and over-use of Nudge

Turning to *classification error* first, any choice to engage in nudging and to require changes to the choice architecture demands a value judgement. It implies a view as to which particular behaviours are seen as ill-informed mistakes or as informed expressions of preferences. This matters, for example, when it comes to the 'selling' of financial products to individuals who are given little insight into the actual risks involved. Furthermore, choosing which products and consumers deserve a 'nudge' in order to ensure that they are not pursuing ill-intended mistakes is a choice as to what one considers to be a 'problem' warranting intervention. A choice between a reliance on nudging or on more prohibitive approaches requires a judgement as to whether the identified problem can be contained to the individual decision maker or whether any failing will impact on the overall trust in markets. Making a choice that certain products do not deserve a 'nudge' to guide individual decision making assumes, firstly, informed decision making and, secondly, that losses are not going to have wider psychological effects on other individuals.

Aggravation suggests that despite nudging the actual problem is getting worse as individuals are provoked into opposing the intended effect. Individuals may regard any form of manipulation as an illegitimate interference in their choices.[3] They may therefore opt out of public provision and thereby increase the actual problem. For example, 'nudging' pupils to eat healthier meals may lead parents to respond by giving their children more of their own food, thereby enhancing obesity. Similarly, forcing companies to remove sweets from the till area may only lead to more colourful and manipulative advertising. A behavioural response that aggravates the problem rather than mitigates it may in particular occur in those situations where the intended outcome is not seen as desirable by at least a minority, especially when this minority is ill intentioned, i.e. hostile to the policy intent, rather than ill informed.

Placation suggests that nudging may lead to some change in behaviour that is then seen as addressing the problem. In fact, however, the long-term problem is not addressed, leading to a much worse problem later. Individuals after being nudged to wear cycling helmets may bike in a more risky fashion as they feel 'safe'; similarly, individuals may eat particular foodstuffs because they are advertised (nudged) as 'healthy' ('green') if eaten in 'reasonable quantities'. As, however, there is no knowledge of what reasonable quantity means, such traffic light labelling can lead individuals to over-consume certain foods that appear healthy, but are not, while they only eat 'good' food in small doses because of some optical signal on the packaging that suggests that over-consumption may not be particularly healthy.[4] Furthermore, being publicly nudged in terms of pension provision may also give a signal that future retirement earnings are safe and at an appropriate level, when they may not be. Individuals therefore may show little interest in considering private savings options. In short, nudging relies on a signal that is 'easy' to understand; it

[3] Indeed, 'nudging' letters by the UK's Revenue and Customs regarding lower than expected tax payments were accused of representing 'bullying' by tabloid newspapers (see Dunn 2014).
[4] This argument was made by the food industry to prevent the introduction of a standardised 'traffic light' system.

may have the unintended effect of individuals caring even less about long-term consequences of their short-term actions.

Over-commitment emerges when Nudge-type initiatives are announced and adopted with much enthusiasm, but then lead to disappointment. One example of over-commitment is where the intended response overwhelms existing administrative capacities. As a consequence, support may decline and therefore increased use of the 'opt out' option will occur. For example, an 'opt out' organ donation system may be widely supported, but will lose support if body parts are found to be distributed through a system of medical favouritism. Another example is where the individual gets overwhelmed by the multiplicity of nudging devices. Similarly, worded statements to incentivise on-time tax payment, multiple labels to inform about food choices and online billing information may lead to a nudging overload that causes a reluctance to be guided by such information (especially with online billing when this requires the retrieval of passwords). A third example of over-commitment is the long-term effect of Nudge. It remains to be seen whether the long-term effectiveness of nudging will wear off, therefore requiring even fancier devices to attract individuals' attention.

A final example of over-commitment and classification error is where compliance with the Nudge does not seem to be forthcoming, for example, when consumers appear to be actively rejecting the set default option. For instance, UK consumers were actively opting-out of government imposed internet filters ('parental controls') in order to be able to watch pornography and other adult websites regardless of the government requiring internet service providers to implement a default setting that blocked legal pornography and other 'adult subjects'. Some blamed this on the engineers who set up private household internet connections rather than actively choosing customers. Regardless, a pattern of 4-8 per cent (for one operator it was approximately 36 per cent) of customers accepting the default setting did suggest that consumers were actively 'opting in' (Miller 2014).[5]

These four mechanisms are not meant to suggest that all Nudge-type interventions are going to fail. It might be argued that all of the issues considered in this section can be ironed out through careful design, i.e. through the application of high degrees of rationality. However, it is unlikely that Nudge can be removed from the context of bounded rationality. Not all nudging will be ineffective, similarly, alternative strategies may be more (or less) useful than Nudge. However, whether Nudge and other behavioural 'insights' are useful remains difficult to assess (but see, Hallsworth et al. 2014; Haynes et al. 2013). For example, the UK House of Lords' Science and Technology Committee (2014) noted how difficult it was, based on published information, to establish which activities of the Behavioural Insights Team had proven effective or ineffective, or why certain initiatives had been

[5] This example has also elements of classification error.

pursued and others had not.[6] The charge that the proof of Nudge was in the pudding (i.e. evidence) rather in its basic assumptions is difficult to assess as the evidence, such as there is, has been produced by those with a stake in Nudge's success. What this section suggests is that a policy tool which claims to be rationalising 'bounded rationality' is faced by limitations introduced by the presence of bounded rationality itself.

Implications

Nudge represents the latest incarnation of 'rational' policy tools to overcome perceived inferior outcomes due to bounded rationality. As noted, Nudge is a somewhat different policy tool as it focuses on the target of the policy intervention, not on the way in which decisions within government are made. However, as previous studies of such 'rationalising' instruments have argued (Wildavsky 1966), Nudge is unlikely to overcome those inherent limitations that affect all government decision making. Nudge has very high rationality assumptions – it assumes the possibility of expert judgement, the possibility to predict the effect of 'architecture' choices, and the possibility of well intentioned individuals' willingness to choose on the basis of being better informed. Nudge is, however, not just like any other tool as it claims to be addressing bounded rationality.

One objection to this argument is that all these claims are largely speculation and that the actual track record is likely to show that Nudge offers effective solutions. Some results may exist that point to such an outcome. Others may argue that the results are hardly insightful as they are based on limited samples, do not rely on a true natural experiment which focuses on different intervention theories, and rely on creatively constructed data. Most of the evidence is produced by actors that have a stake in the advocacy of Nudge. In other words, it is difficult to argue that straightforward evidence in favour or against Nudge can be produced.

Nudge suffers from an over-optimism in its rationality assumptions that is, as noted, particularly surprising in that it addresses bounded rationality directly. Nudge is not sufficiently reflective of its own limitations. As such, this limitation may also not be unexpected as ideas about benevolent governments and 'imperfect' private decision making are not unusual for public policy writing: Nudge and 'behavioural public policy' seem to be affected by the academic biases of their advocates.

[6] The Behavioural Insights Team responded by suggesting that much of its work was in the peer review process for academic journals, but pointed to its 'EAST' publication (Behavioural Insights Team 2014) for evidence of examples that presented 'what we have learnt about what works in this field – and (importantly) what does not'. That 'EAST' report did contain examples of experiments where interventions had shown effects and others had not, without being able to offer an explanation for different effects apart from stressing specificity and personalisation. The report also included 'pitfalls', which either drew on earlier literature or offered examples where the Behavioural Insight Team's work was not found to have failed to detect noticeable results in at least one intervention strategy.

The direct implication therefore is that Nudge enthusiasts in government (and beyond) should be nudged into considering further the limits of their knowledge, and the type of experimental evidence base they are relying on and the inherent trade-offs and side effects that occur in organisational and individual decision making. Wildavsky (1983) argued that one of the key limitations of the policy sciences was their lack of consideration of the importance of organisational logics within government. It is indeed highly ironic that bounded rationality is used to bring rationality to policy, without acknowledging the conditions of bounded rationality. This does not mean that all government intervention should be discarded, or that governments are incapable of advancing options that support individuals in ways that these targeted individuals could not pursue themselves. Instead, what is required is somewhat more modesty when it comes to the pursuit of 'rational' policy making in executive government.

Given the limitations of Nudge and other behavioural tools, attention naturally shifts to the sources of these tools' contemporary popularity. The persuasive appeal of Nudge (Hood & Jackson 1991,1994; Majone 1989) points to the conditions that explain the current policy boom (Dunleavy 1986). Nudge appeals to diverse constituencies, whether it is those that emphasise the choice element, the possibility of using evidence, and to shape private choice for public gain. It appeals to academic fashions, especially those linked to the wider hegemony of quantitative policy science. It also appeals politically as it suggests 'cheap government': governing through nodality rather than through treasure, authority or organisation offers a promise of little resource depletion (Hood & Margetts 2007). While every little cost saving helps, it is questionable whether governing through nodality can be done on the cheap as depleted states are unable to address policy challenges in more resourceful ways. It is also questionable how far Nudge can be taken, as it assumes well intentioned and ill informed individuals. It may therefore have little to say about suicide bombers, corporate tax avoiders, or other 'amoral calculators'. Similarly, it is doubtful whether Nudge is applicable to the ill informed and ill intentioned, such as estate agents.

In conclusion, questions of individuals and organisations' motivations and capacities are clearly critical for the practice and study of politics and public policy. This paper has argued that while the limits of Nudge resemble those of other temporary policy fashions and movements, it is particularly noteworthy and ironic that an approach claiming to be dealing with reasoning failure and bounded rationality shows so little self-awareness of its own limitations.

References

6, P. (2014) 'Explaining decision-making in government', *Public Administration* 92 (1): 87–103.

Abadie, A. & Gay, S. (2006) 'The impact of presumed consent legislation on cadaveric organ donation: a cross-country study', *Journal of Health Economics* 25 (4): 599–620.

Amir, O. & Lobel, O. (2008) 'Stumble, predict, nudge: how behavioral economics informs law and policy', *Columbia Law Review* 108 (8): 2098–2137.

Ariely, D. (2008) *Predictably irrational: the hidden forces that shape our decisions,* New York: Harper Perennial.

Baldwin. R. (2010) 'Better regulation: the search and the struggle', in R. Baldwin, M. Cave and M. Lodge (eds), *Oxford Handbook of Regulation,* Oxford: Oxford University Press.

Baumgartner, F. and Jones, B.D. (1993) *Agendas and instability in American politics,* Chicago: University of Chicago Press.

Behavioural Insights Team (2014) 'EAST – Four simple ways of to apply behavioural insights', London: Behavioural Insights Team. <http://www.behaviouralinsights.co.uk/sites/default/files/BIT%20Publication%20EAST_FA_WEB.pdf> Accessed 19 November 2014.

Choi, J.J., Laibson, D., Madrian, B.C., & Metrick, A. (2001) 'Defined contribution pensions: plan rules, participant decisions, and the path of least resistance', NBER Working Paper No. 8655, Cambridge MA: National Bureau of Economic Research.<http://www.nber.org/papers/w8655.pdf>

Cohen, M., March, J. and Olsen, J.P. (1972) 'A garbage can model of organizational choice', *Administrative Science Quarterly* 17: 1–25.

Dunleavy, P. (1986) 'Explaining the privatization boom: public choice versus radical approaches', *Public Administration* 64 (1): 13–34.

Dunn, S. (2014) '"They should stop picking on hard-working people": taxman's bully-boy letters to innocent families', *Daily Mail,* 30 May. <http://www.dailymail.co.uk/money/news/article-2644274/Taxmans-bully-boy-letters-innocent-families.html> Accessed 24 June 2014.

Gigerenzer, G. (2014) *Risk savvy,* London: Allen Lane.

Hallsworth, M., List, J., Metcalfe, R. and Vlaev, I. (2014) 'The behavioralist as tax collector', NBER Working Paper no. 20007, Cambridge MA: National Bureau of Economic Research. <http://www.nber.org/papers/w20007> Accessed 2 August 2014.

Harford, T (2014) 'Big Data: are we making a big mistake?', *Financial Times,* 28 March. <http://www.ft.com/cms/s/2/21a6e7d8-b479-11e3-a09a-00144feabdc0.html#axzz3Ab4bTQHF> Accessed 16 August 2014.

Haynes, L., Green, D., Gallagher, R., John, P. and Torgerson, D. (2013) 'Collection of delinquent fines', *Journal of Policy Analysis and Management* 32 (4): 718–30.

Hood, C. (1976) *The limits of administration,* London: Wiley.

Hood, C. & Jackson, M. (1991) *Administrative argument,* Aldershot: Dartmouth Publishing.

Hood, C. & Jackson, M. (1994) 'Keys for locks in administrative argument', *Administration & Society* 25(4): 467–88.

Hood, C. & Lodge, M (2005) ' Pavlovian innovation, pet solutions and economizing on rationality', in J. Black, M. Lodge & M. Thatcher (eds), *Regulatory intervention,* Cheltenham: Edward Elgar.

Hood, C. & Margetts, H. (2007) *Tools of government in the digital age,* Basingstoke: Palgrave.

House of Lords (2014) Science and Technology Select Committee Letter to Rt Hon Oliver Letwin MP, Minister for Government Policy and Chancellor of the Duchy of Lancaster, Cabinet Office, 22 July. <http://www.parliament.uk/documents/lords-committees/science-technology/behaviourchangefollowup/LetwinBehaviourChangeLtr20140722.pdf> Accessed 2 August 2014.

John, P. (2011) *Making policy work,* London: Routledge.

John, P., Cotterill, S., Moseley, A., Richardson, L., Smith, G., Stoker, G., & Wales, C. (2011) *Nudge, nudge, think, think: experimenting with ways to change civic behaviour.* New York: Bloomsbury Academic.

Jones, B.D. (2003) 'Bounded rationality and political science: lessons from public administration and public policy', *Journal of Public Administration Research and Theory* 13 (4): 395–412.

Kahneman, D. (2011) *Thinking, fast and slow,* New York: Farrar, Straus and Giroux.

Kahneman, D. & Tversky, A. (1979) 'Prospect theory: an analysis of decision under risk', *Econometrica* 47(2): 263–91.

King, A. & Crewe, I. (2013) The blunders of our governments, London: Oneworld.

Kingston, A. (2012) 'Medicine's deadly gender gap', *Maclean's,* 30 April. <http://www.macleans.ca/society/health/medicines-deadly-gender-gap-2/> Accessed 1 September 2013.

Lindblom, C. (1959) 'The science of "muddling through"', *Public Administration Review* 19 (2): 79–88.

Lodge, M. & Wegrich, K. (2014) 'Crowdsourcing and regulatory reviews: a new way of challenging red tape in British government?', *Regulation & Governance*, doi: 10.1111/rego.12048 .

Madrian, B.C. & Shea, D.F. (2000) 'The power of suggestion: inertia in 401(k) participation and savings behavior', Working Paper no. 7682, Cambridge MA: National Bureau of Economic Research. <http://www.nber.org/papers/w7682> l Accessed 18 November 2014.

Majone, G. (1989) *Policy, argument and persuasion*, New Haven CT: Yale University Press.

March, J.G. and Olsen, J.P. (1983) 'The new institutionalism: organizational factors in political life', *American Political Science Review* 78 (3): 734–49.

March, J.G., Sproul, L. and Tamuz, M. (1991) 'Learning from samples of one or fewer', *Organization Science* 2 (1): 1–13.

Merton, R (1936) 'The unanticipated consequences of purposes social action', *American Sociological Review* 1 (6): 894–904.

Miller, G. (2000) 'Rational choice and dysfunctional institutions', *Governance* 13 (4): 535–47.

Miller, J. (2014) 'New broadband users shun UK porn filters, Ofcom finds', BBC, 23 July. <http://www.bbc.co.uk/news/technology-28440067> Accessed 19 November 2014.

OECD (2010) *Obesity and the economics of prevention*, Paris: OECD.
Oliver, A. (2013) 'From nudging to budging: using behavioural economics to inform public sector policy', *Journal of Social Policy* 42 (4): 685–700.

Pierson, P. (2000) 'The limits of design: explaining institutional origins and change', *Governance* 13 (4): 475–99.

Rizzo, M.J. & Whitman, D.G. (2008) 'The knowledge problem of new paternalism', *Brigham Young University Law Review* 4: 905–68.

Rizzo, M. & Whitman, D. (2009) 'Little Brother is watching you: new paternalism on the slippery slopes', SSRN Scholarly Paper no. 1119325. Rochester NY: Social Science Research Network <http://papers.ssrn.com/abstract=1119325> Accessed 19 November 2014.

Rose, R. (1990) 'Inheritance before choice in public policy', *Journal of Theoretical Politics* 2 (3): 263–91.

Sieber, S. (1981) *Fatal remedies*, New York: Plenum Press.

Simon, H.A. (1997) *Administrative behavior: a study of decision-making processes in administrative organization*, 4th edn., New York: Free Press. First published 1945.

Stark, A. (2014) 'Bureaucratic values and resilience: an exploration of crisis management adaptation', *Public Administration* 92 (3): 692–706.

Thaler, R.H. & Sunstein, C.R. (2008) *Nudge*, New Haven CT: Yale University Press.

Wildavsky, A. (1983) 'Information as an organizational problem', *Journal of Management Studies* 20 (1): 29–40.

Wildavsky, A. (1966) 'The political economy of efficiency: cost-benefit analysis, systems analysis, and program budgeting', *Public Administration Review* 26 (4): 292–310.

Wilkinson, T.M. (2013) 'Nudging and manipulation', *Political Studies* 61: 341–55.

Centre for Analysis of Risk and Regulation

The London School of Economics and Political Science
Houghton Street
London WC2A 2AE

tel: +44 (0)20 7955 6577
fax: +44 (0)20 7955 7420
email: risk@lse.ac.uk

lse.ac.uk/CARR