
24. Behaviour change: extralegal, apolitical, scientific?

Robert Lepenies and Magdalena Małecka

1. INTRODUCTION

Nudges pose challenges for democracy and the law and they presume a simplified view of science. These very broad charges will be sustained through a series of sketched arguments that criticize the intellectual foundation, political implementation, and democratic compatibility of behavioural public policy. Our three critiques differ from frequent criticism of nudging in that we are less interested in establishing whether a certain application of behavioural insights is manipulative, or violates the values of welfarism, autonomy or dignity. Rather, we analyse whether the possibly widespread adoption of behaviour change undermines broader institutions which society values.

Recently, Cass Sunstein has responded to critics of behavioural change by warning them about falling into the ‘trap of abstraction’ (Sunstein 2015). Here, we would like to give three examples of how abstract analysis can add to our understanding of behaviour change, highlight problematic tensions, and how only through looking at the institutional implications of behaviour change, we can search for safeguards of societal institutions.

2. RESCUING LAW FROM BEHAVIOUR CHANGE?

In what sense could one make the claim that the law has to be rescued from the increasing reliance upon behavioural insights? How is law in tension with behaviour change? We argue here that there is a sub-set of behavioural policy instruments which undermines the institution of the law. Put simply, behavioural instruments can undermine legal norms and thereby hinder societies to engage in what we call collective self-legislation as they are publicly invisible.¹

We propose an institutionalist perspective that looks at broader societal implications of behaviour change, rather than a mere individualist perspective which is common in the literature (as we show in Lepenies and Małecka 2015). To better understand the impact that behavioural policy has on the law, we summarize and expand on prior work (Lepenies and Małecka 2015) by comparing two ideal types of how the state can shape behaviour in societies: the normative and the instrumental way. For the former, law must influence behaviour through legal norms. For the latter, effective behaviour change does not require cognitive awareness of instrument that impacts behaviour.

Our perspective is at odds with the growing enthusiasm displayed by many legal scholars about behaviour change. In what follows, we will outline our critical institutional perspective and compare it to other evaluations in legal studies and legal theory before suggesting some safeguards that can make behavioural instruments compatible with legal norms.

3. NUDGES AND BEHAVIOURAL INSTRUMENTS

Nudges are intentionally designed policy instruments which use insights from the behavioural sciences to steer people's behaviour. For this, they rely on empirical insights about human behavioural regularities. As a sub-set of behavioural instruments, nudges represent one behavioural policy tool. Importantly, however, nudges do not always change behaviour. Indeed, the most powerful types of nudge, default rules, rely on the tendency of humans to be inert and accept pre-defined settings. Proponents of nudging argue that they gently coax people into a particular action – that they are welfare-promoting yet choice-preserving, while allegedly not limiting individual freedom of choice. Nudges are intended to impact behaviour in a mode distinct from rational persuasion, command-and-control instruments, or (material) incentives. Different from prior work (Lepenies and Małecka 2015), we here focus not only on nudges, and if we talk about nudging, we deal with only certain kinds of nudges. As Sunstein (2015, p. 525) rightly responded to us: 'The defining feature of nudges is not that they are invisible, but that they preserve freedom of choice.'

Instead, we aim to make our prior account more precise by focusing only on behavioural instruments that are non-normative and non-cognitive: is the behaviour of citizens shaped by drawing on their understanding of legal norms, or are these bypassed through the help of knowledge about behavioural regularities? We argue that when looking at the impact of behavioural instruments such as nudges, we should not look only at what a nudge does or does not do to a nudgee, but also whether the widespread adoption of nudging as a policy has institutional implications (owing to the nature of the instrument in question). Due to their idiosyncratic intellectual genesis and due to the lack of one widely accepted definition, our account looks at nudging only in cases where it is non-cognitive and non-normative.

3.1 Behaviour Change Instruments and the Law

There are two ideal types of influencing behaviour through policy and the law: 'the normative way of influencing behaviour', and 'the instrumental way of influencing behaviour'. We identify legal norms with the first type of influence. The distinction helps us to understand and evaluate the way in which nudges and other behavioural instruments shape behaviour and the way in which public rules are understood by policy-makers and legislators. In effect, the two ways of influencing behaviour correspond to two different ideal types of guiding behaviour in the public sphere: through normative – or through non-normative ways.

The normative way of influencing behaviour has the following features: First, there exists a requirement to behave in a specific way. Second, agents must recognize this requirement. Indeed, they must treat the requirement as a reason for action. Third, the requirement needs to be cognitively accessible. This is a necessary condition for effective responding to the influence. The 'instrumental way of influencing behaviour' has the following features: First, there is a change of the context of individual decision-making. Second, agents react to the changes in the choice architecture (that affect, for instance, perception of risk, emotions). Third, the choice architecture does not need to be cognitively accessible in order to be effective. In other words, the non-cognitive reaction of an

agent is a sufficient condition for effective responding to the influence. We argue that the first ideal type approximates the way in which behaviour is influenced by legal norms, whereas the second ideal type approximates the way in which behaviour is influenced by behavioural instruments, as they are understood by the proponents of bringing evidence from the behavioural sciences to policy. Certain nudges influence behaviour in a non-normative way, when they do not operate, contrary to influences by legal norms, through requirements and hence do not provide a reason for action. But even those behavioural instruments that are part of a legal system, such as opt-out rules, affect situations of agents without giving reasons for their actions.

Through knowledge about human behavioural tendencies, policy-makers are able to influence behaviour in a non-cognitive way. By this we do not mean to say that these behavioural instruments work by excluding all cognitive effort by citizens. But they are non-cognitive in the sense that citizens are not aware of why and when cognitive effort is part of the desired behaviour (which choice architects or ‘planners’ stipulate). To be effective, it is sufficient that people react correctly (in the eye of the ‘planner’²) to the choice architecture in question. This contrasts with the normative function of the law. Here, the cognitive accessibility to the content of a legal norm (requirement) is a necessary condition without which there can be no impact on behaviour. Thus law, in order to have this impact, has to be publicly accessible and visible.

These two ways of influencing people’s behaviour correspond to two visions of agency. Behaviour change holds a view of human agents in which individuals act out of emotions, sentiments, automatic reactions. In the normative view, on the other hand (note that we are speaking about ideal types here), people deliberate about norms, discuss whether they (should) bind them, and, thus, whether they act as sufficient reasons to obey norms and accept the stipulated requirements.

Sunstein (2014) admits that nudges might be ethically troubling in case they rely on unconscious cognitive processes, or emotions, or both (the so-called System 1). He argues that since in our decision-making, we cannot avoid being influenced by System 1 processes, therefore ‘so long as the initiatives are made public and defended on their merits, nudges should not be ruled off-limits merely because they work as a result of the operations of System 1’ (Sunstein 2014, p. 151).

In the behaviour change agenda, evidence for the effective impact of a behavioural intervention – whether drawing on so-called System 1 or System 2 – is provided through experimental methods in policy trials, and almost exclusively through using randomized controlled trials (RCTs). Proof of effectiveness (of impacting behaviour) drives implementation; yet citizens have no way of knowing this. The point here is not to fault RCTs in terms of research ethics. Rather, when RCTs are determining policy instrument selection, they enter as non-participatory and non-deliberative components of the policy process. Through this, behavioural instruments that are selected upon an experimental basis are much more difficult to make transparent than, for example, the simple distinction between System 1 and System 2 nudges suggests. Even if behavioural instruments are disclosed, the mechanisms by which they are chosen withstands such disclosure.³

In contrast, the law influences behaviour through publicly accessible rules which appeal to reason in a non-instrumental way. Legal norms provide an important institutional role: the law enables collective self-legislation. We understand self-legislation as the control a social collective has over its evaluation, deliberation and choice of social institutions (this

definition is inspired by a related definition of autonomy by Hausman and Welch 2010). Here, law that approximates ‘the normative way of influencing behaviour’ may fulfil a dual function of motivating individual behaviour, as well as at providing a social ideal towards which action should strive.

A similar view has also been provided by Peter Cserne who proposes to understand the law as giving normative guidance, ‘according to which the law gives reasons for action to its addressees who should be able to either obey or disobey the law’ (Cserne 2015, p.293). Cserne points out that this understanding ‘is widely shared by the legal and political community and that it presumes a certain vision on human agency as triggered by motives and reasons’. Thus, in both Cserne’s and our account,⁴ drawing on legal norms when impacting behaviour also means endorsing a view of agency that is compatible with democratic deliberation over the rules that society gives itself. Elsewhere, Cserne notes, drawing on Fuller (1964, p. 162) that the law ‘involves of necessity a commitment to the view of man that is, or can become, a responsible agent, capable of understanding and following rules, and answerable for his defaults’ (Cserne 2016, p. 167).

In which sense can these behavioural instruments be extra-legal? We argue that nudges do not impact behaviour in the way legal norms do. As they are typically introduced not through parliamentary legislation, but by acts of governmental and administrative bodies, it is worth asking whether, or under which conditions they remain legal.⁵

Cserne (2016) distinguishes different types of legality and argues that depending on one’s view of the law, behavioural policy, even if formally authorized, might be extra-legal or even pervert the law. In his view, however nudges are not extra-legal ‘in an doctrinal sense’. Despite being forms of ‘techno-regulation’, there are ‘legal principles and doctrinal techniques such as human rights, the rule of law or proportionality ... available for the task of control’ (Cserne 2016, p.170). And indeed, other legal scholars have discussed precisely these issues of legal control, for example, on the constitutionality of nudging (Kemmerer et al. 2016; Alemanno and Sibony 2015). Anne van Aaken (2016, see also Chapter 21 in this handbook) has enquired whether nudges might be held to run afoul of specific legal requirements, looking at German constitutional law. She doubts that many nudges will satisfy legal tests of proportionality that are standardly used (van Aaken 2016, p. 136) and argues that nudges, as an instance of soft law, are difficult to control and contest legally. McCrudden and King (2016) similarly raise broader questions about how behavioural instruments impact the legal system. While they note that increasing discretion of public agencies to use behavioural insights might skew legal checks and balances, they are confident about the ability of courts to effectively police this. Their worry pertains more to nudges not being part of formal rule-making procedures, and the threat of nudges displacing traditional formal democratic rule-making in case they become the default option. They point to the 2011 House of Lords Science and Technology Select Committee Behaviour Change report saying that nudging diverts government from its responsibility to use other more effective instruments (McCrudden and King 2016, p. 92).

3.2 Safeguards

The legal system – as the body of codified, publicly accessible and debated legal norms – is a social institution which can make nudges more visible and accessible and thus minimize a non-normative and a non-cognitive impact of nudges on citizens’ behaviour. Here,

we propose safeguards which should relate nudges to the legal system in a visible, and recognizable way. This connection of nudges with the legal system should make agents reflect on the influences they are subject to, as well as being able to debate, and in principle – oppose, the (legal) preconditions of such influences.

Despite the tensions between the ideal types of ways of influencing behaviour, we nevertheless want to explore how law could act as a safeguard for possible negative consequences of behavioural instruments. We present a rather loose list of suggestions in order to stimulate debate on policy solutions. These include wider liability rules, behavioural disclosure labelling, expiration dates, nudge registries, and novel judicial and legislative review mechanisms that ensure that nudges do not replace legal norms but are instead complementary, or a ‘nudging ombudsman’. At the moment, there are few safeguards to nudges in place and few discussions about their institutional consequences.

One suggestion to address potential abuse of behavioural instruments (and default rules in particular) is to strengthen liability stipulations for lawmakers in the case that interventions infringe rights or otherwise violate the constitutional order. Those decisions might be taken by ‘nudging oversight bodies’, for example, a quasi-public, quasi-independent ‘nudging ombudsman’ which could be appointed by parliaments. The task of such a body would be to represent a broad variety of societal and legal concerns, as well as diverse academic and scientific perspectives. It would oversee the conformity of nudges generally (not just default rules) to constitutional and basic legal principles (see Alemanno and Spina 2014 for a more limited proposal to introduce regulatory impact assessments; we are sympathetic to their idea to expand the scope of judicial review when assessing nudges).

A further helpful suggestion has been made by McCrudden and King (2016) who advocate for clearer boundaries of administrative agencies when employing behavioural instruments to not overstep their respective areas of influence. This opens up the more general case for the need of regulating nudge units themselves, a proposal that has not yet been made in the literature.

Another proposal would be behavioural disclosure labelling: A good example for this are shocking traffic warnings (e.g. billboards on highways that show crashed cars or disfigured bodies). These warnings are not part of the legal system in the narrow sense of having been codified and formalized in law. They supposedly work through drivers’ reliance on availability heuristics. Drivers effectively react to the warning on the billboard (and drive slower), without responding to the law or a legal norm. Interventions that have been experimentally tested as being more effective than alternative policy proposals are publicly invisible in the sense described above. They are, in this case, also relying on System 1 processes, and cannot be easily controlled, reacted to, or resisted. We think that in order to avoid misusing this kind of policy solutions, shocking traffic warnings (just like shocking health warnings) should be complemented with information about the legal source of such warnings, as well as by the liability of law or policy-makers who misuse this kind of instrument. Akin to the European Food Safety Agency which is responsible for the labelling of food additives through so-called E-numbers, one tongue-in cheek idea would be to establish a Nudging Safety Agency which likewise labels individual interventions.

Probably a more serious contender for bringing behavioural instruments closer to legal norms is to make them more visible through a legal registry of behavioural instruments (a ‘nudge registry’). National or sub-national administrative bodies would have to disclose

type, duration, purpose, and mechanism of behavioural interventions and provide a written justification for why a certain nudge is employed. Such a registry and the accompanying mandatory justifications would direct the debate away from a debate about (cost-) effectiveness and instead point to the broader context in which behavioural instruments are to be implemented: an institutional perspective. Our general claim remains the following: ‘policy-makers ought to try as far as possible to complement nudges with legal norms or to make nudges connect with a legal system (or in other words – bring them into the legal system). The rationale for our proposal is to make nudges accessible and public and in this way weaken their non-cognitive and non-normative impact on agents’ behavior’ (Lepenies and Malecka 2015, p.435).

4. RESCUING POLITICS FROM BEHAVIOUR CHANGE?

Above we have made the case for behavioural change as undermining legal norms. But it is also possible to expand on this analysis by outlining how behavioural policy undermines politics. We treat behaviour change here as a novel form of governing which, if widely adopted, threatens to undermine traditional justificatory relationships between policy-makers and citizens. What does politics look like in a ‘nudge world’?

We will make the case that the behaviour change agenda is in tension with a range of conceptions of politics. Our argument is less about a reconstruction of the explicit content of political justifications given by proponents of behavioural policy. Rather, we highlight that behavioural instruments eliminate reason-giving, or ‘political normativity’ in society.

Proponents of behaviour change have themselves called for the need to debate nudges in public – yet, they do not see the more fundamental tension between behaviour change and political normativity. Sunstein (2016) suggests that there are nudges that ‘can claim a democratic pedigree and that promote democratic goals’. In fact, he argues ‘A self-governing society might well nudge its citizens to participate in the political process and to vote’ (Sunstein 2016, p.14). Nudges that encourage turnout, or automatic voter registration defaults are desirable in his eyes for a healthy polity. We see this endorsement of such ‘democracy nudges’ as problematic, and not compatible with the value of collective self-legislation described above if they are introduced as ‘nudges’ by nudge units, and not introduced as, for example, pieces of legislation discussed and passed by parliaments. With increasing endorsement of behavioural policy, these tensions become more salient.

Much ink has been spilled on describing the political nature of the behaviour change agenda. Famously, *Nudge* (Thaler and Sunstein 2008) provided the first explicitly normative defence of the behaviour change project. Several behavioural insights teams either formally or informally still refer to the concept of nudge, openly endorsing the values of libertarian paternalism (see for example newly founded teams in Japan or in Qatar which call themselves nudge units). With increasing institutionalization of behavioural policy, many practitioners however have begun to dissociate nudging from what they argue is a broader agenda of behaviour change, understood as bringing the behavioural sciences to policy. Most proponents now claim that the application of behavioural instruments is independent of any particular ideological outlook (Lourenço et al. 2016). Instead, it is claimed that behavioural instruments can be applied to achieve a broad range of purposes that go beyond libertarian paternalism. As will be seen in the next section, it is behavioural

science that is invoked as a justification for policies, and not strictly and only behavioural economics. See for instance Hallsworth (2016, p. 42): ‘most policies influence behavior and therefore behavioral science has something to say about most policies’. Behaviour change is bigger than nudging, and it is bigger than behavioural economics. Instead, behaviour change is about bringing (behavioural) science to government. Thaler for instance writes:

The work of the BIT has often been mischaracterized as being based on behavioral economics whereas, in fact, there has been, at least up to now, very little actual economics involved. The tools and insights come primarily from psychology and the other social sciences. The whole point of forming a Behavioural Insights Team is to utilize the findings of other social sciences to augment the usual advice being provided by economists. It is a slur to those other social sciences if people insist on calling any policy-related research some kind of economics. (Thaler 2015, p. 334)

If we agree with Thaler, then most contemporary critics of behaviour change are mistaken in seeing behaviour change as a project of neoliberal governmentality, as their reasoning is based on a critique of behavioural economics. We move to this critique in the next section.

4.1 The Politics of Nudging and Behaviour Change: Critics

Some critics of the nudge approach, like McMahon for instance, have taken a Foucauldian perspective when arguing that behavioural economics is a ‘political economic apparatus of neoliberal governmentality’ (McMahon 2015, p.1). For McMahon, the ‘unequivocal objective of behavioural economics is to cultivate subjects that more closely conform to market logics’ (p. 1). A similar argument is made by Leggett (2014) and Jones et al. (2011). Zuidhof, however, notes in his case study of behavioural applications by the European Commission (Chapter 12 in this handbook) that rather than only being concerned with economic outcomes, ‘behavioral economics may ... surface in areas where more invasive policies run into political limitations’. Behavioural policy therefore only depoliticizes certain policy areas. Similar verdicts on partial or universal depoliticization are brought by McMahon (2015) and Leggett (2014).

We contend that behaviour change cannot be understood by looking at behavioural economics or libertarian paternalism alone. Indeed, a shift can be observed away from the rhetoric and endorsement of nudging and libertarian paternalism, with newer behavioural insights teams utilizing the methods or rhetoric of evidence-based policy (Germany is the best example here, which has ceased to speak of ‘nudging’ in public relations). But as Straßheim and Kettunen (2014) argue, claims concerning evidence-based policy need to be socially contextualized and understood as embedded in political and epistemic struggles.

4.2 Choose Only One: Effectiveness or Politics

There has been some speculation over why it has been the case that most behavioural instruments have not been imposed by lawmakers as codified requirements. One guess is that decisions about implementing behaviour change policies are driven by the effectiveness of the instrument. Behaviour change instruments do not require much political support, parliamentary procedure and debate. In this way, they enable policy-makers to influence people’s behaviour more quickly, more effectively, and without putting effort into legislative and deliberative processes. Behavioural policy is spreading, and can be

found in governments around the world (OECD 2017). We see this as an empowering of the executive branch of government as new effective tools of influencing behaviour can be employed, with fewer strings attached (we do not know whether this empowering is cause or symptom of a more general strengthening of the executive). In turn, as behavioural instruments can be employed by different parts of public administration, for example, by several sub-national agencies, they can spread more easily than legal norms, whose formulation, implementation and revision is often cumbersome. It is also specific agencies that, due to the informal nature of most nudges, have found to have quite some lee-way in applying nudges (see also McCrudden and King 2016). Much of the heterogeneity between units is unexplored. Yet prominent proponents of behavioural policy have themselves described their approach as radical incrementalist (Halpern and Mason 2015) and seem to share a vision of political change, improving the world ‘one RCT at a time’.

In the past, much optimism was connected to the ability of behavioural policy to overcome partisan divide in policy as ‘the real third way’ (Thaler and Sunstein 2008, p. 252) and of nudges as charting a ‘viable middle ground in our unnecessarily polarized society’ (Thaler and Sunstein 2008, p. 252). Today however, nudge units have themselves become the object of political economy: increasing in size, they are now political agents themselves that are influenced by the political context they are operating in (Schubert 2017).

4.3 The Disappearance of Political Normativity through Invisible Instruments

Critical remarks emphasizing the threat of nudges to policy visibility, institutional transparency, and public deliberation have already been formulated by some scholars (Hansen and Jespersen 2013; John et al. 2009). Sunstein, in his response to this kind of critique, remarks that, since nudges are often highly salient, they are by nature transparent and debated in public (Sunstein 2014), even though nudges that rely on unconscious cognitive processes might at times be problematic.

As early as 2008, proponents of behaviour change emphasized that nudges need to be public, using a publicity principle as one of their ‘guiding principles’ (Thaler and Sunstein 2008, p. 244) and as ‘a good guideline for constraining and implementing nudges’ (Thaler and Sunstein 2008, p. 245) invoking John Rawls. With Rawls, we agree that rationally autonomous individuals are owed justification for the laws they are subjected to. Citizens ‘must have knowledge of the moral bases of coercive laws’, and of the ‘real reasons for their social and political relations’; it is publicity of the first principles which should give ‘democratic citizens a common basis for political argument and justification’ (Freeman 2007, p. 187). As all codified rules, these are in principle publicly accessible and thus visible in our sense.

There is a long tradition in democratic theory that sees the legitimacy of political decisions generally dependent upon the procedures and quality of deliberation by which they were reached. In this vein, John et al. (2009) have very early juxtaposed ‘nudge approaches’ to ‘think approaches’, making a very similar claim to us, even calling for ‘more attention to the way that collective and constitutional settings help determine the success or failure of a nudge’ (John et al. 2009, p. 369). Their Think approach starts from insights from the deliberative turn in democratic theory, where public deliberation is at the core of politics. This centrality of public justification is shared by a range of democratic theorists from Habermas to Forst.

We argue, however, that behavioural policy crowds out such deliberative engagement with invisible policy tools. Sunstein replies here that '[i]nsofar as they come from government, nudges need not (and should not) suffer from any problem of invisibility' and further, '[e]verything that government does should be scrutinized' (Sunstein 2014, p. 584). Indeed, 'Any official nudging should be transparent and open rather than hidden and covert. Indeed, transparency should be built into the basic practice' (p. 584). But what if this is not possible?⁶

For our account, it is not decisive whether individuals understand the intervention itself (e.g. citizens understand that a smiley face aims to encourage behaviour), but rather, that individuals are not aware that there has been a behavioural/experimental trial which concluded that a particular intervention (the smiley face) would be the most effective way to change their behaviour.⁷ It is ignorance of the origin and full working of the instrument (which includes instrument choice) that we think has problematic consequences.

4.4 Diachronic Effects

One last problematic implication for political institutions is the diachronic effect of non-cognitive and non-normative behavioural policy instruments. We contend that due to their collective invisibility (as outlined above), it is increasingly problematic if societies 'inherit' nudges from prior public administrations. This is a problem that some countries begin to face after the first wave of nudge units (e.g. the UK BIT becoming private; the US SBST being established). Once installed, and by their very nature, nudges are impossible to reverse, as they cannot be easily detected and there is no formal way to dial them back. Behavioural instruments may in this sense unduly limit the collective capacity to shape public rules. Similar to other instances of technocratic governance, behavioural instruments tend to stifle democratic deliberation about the means and ends of policy.

4.5 Safeguards

Our proposed safeguards are similar to those suggested above: policy-makers ought to try as far as possible to re-politicize behavioural instruments in order to connect with the political system. The rationale for our proposal is to make nudges accessible, capable of being argued against, and rendered public. In this way, their non-cognitive and non-normative impact on agents' behaviour might be weakened.

How can public deliberation about the means and ends of behavioural policy be safeguarded? How can behaviour change be brought into the political sphere? Sunstein (2016) agrees with the potential of having 'advance notice' and 'public comment' on behavioural interventions in a 'focused process of public scrutiny and review' which will make sure that public officials will hear about problematic aspects of certain interventions.

Our proposals go further yet: We propose expiration dates for nudges which would force lawmakers to deliberate on implementation in regular intervals. Further, more deliberative and participatory formats should be explored (as has been explored, for example, in behavioural policy in France). Rather than basing the legitimacy of interventions just on surveys of behavioural policy (see also Sunstein in this volume), citizens should be given the chance to reason about behaviour change, ideally through public fora, citizen conferences and townhall meetings.

Likewise, nudge units should increasingly accept their status as political agents and think about which ownership structures and internal decision-making structures are appropriate and legitimate. Another challenge for nudge units is to maintain academic, ideological and political diversity as their influence grows. One possible way to do this is through increasing engagement with political scientists and organizational theorists: nudge units can profit enormously from the insights of these fields, and to compare themselves to similar movements. This will be helpful not just to learn about the success conditions for political movements, but also to become more self-aware about their role as political actors.

5. RESCUING SCIENCE FROM BEHAVIOUR CHANGE?⁸

Behavioural interventions such as nudges are policy instruments defended and promoted as ways of influencing citizens' behaviour on the basis of scientific knowledge (the 'behavioural sciences'). It is believed that this allows for proposing policies that are effective, that 'work', as they are based on the knowledge of what are the regularities of people's behaviour. The idea of using knowledge about the regularities and tendencies of human behaviour to make policies effective (i.e. to bring about desired effects) presumes a rather simplified view of science (in this case – the behavioural sciences) as a repository of facts, or findings that are then being applied in the practical context (see the prominent examples of this view: Jolls, Sunstein and Thaler 1998; Shafir 2012; Sunstein 1997). It presumes that one can clearly distinguish between the scientific endeavour and its application to the policy context. However, this poses philosophical problems of scientific overreach (Boudry and Pigliucci 2017; Pigliucci 2015) and raises problems with the integrity of scientific testimony of those engaged in behaviour change (Anderson 2011). Ultimately, these might impact the integrity of both policy and science.

The simplified view on science is problematic because it obscures the process of knowledge production in science as completely value-free and it presents the policy-making as based on uncontested unquestionable evidence. As such the behaviour change agenda exhibits features of scientism. The plea to 'bring in the behavioural sciences' generally gives scientific justifications with which choice architects could venture into new policy fields and applications. The problem here is one of scientific overreach.

We point out two overlooked features of contemporary science, and of behavioural science more specifically, that challenge this view and pose challenges for the behaviour change agenda. First, we discuss the ways in which values enter scientific processes. Second, we emphasize the consequences of the fact that the behavioural sciences do not form an integrated body of knowledge and that therefore the reliance on behavioural findings in policy-making is necessarily selective. Looking at how nudge units have operated, it was in fact a narrow slice of psychology and economics (and of the behavioural sciences) that informed policy. This selectivity poses considerable long term risk not just for the behaviour change project, but for the influence of science on policy in general without proper safeguards.

In contemporary philosophy of science there is widely accepted consensus that so-called non-epistemic values – political, social and ethical ones – cannot be separated from the processes of knowledge production. Values are an integral part of scientific research

(Douglas 2009; Longino 1990; Kourany 2010; Wylie and Nelson 2007; Solomon 2001). Once we understand the role of values in scientific research, we should see that values are entering the very process of producing knowledge within the behavioural sciences; they do not appear only at the stage of applying this knowledge outside science, that is, in policy. Philosophers point out that non-epistemic values (normative and emotive commitments that concern moral and social life) can influence the choice of topics and of goals that research is expected to serve. Values can influence choices about the acceptance of hypotheses and theories, for example, when scientists have to decide whether the evidence is sufficient to support a claim or hypothesis. Often non-epistemic values are relied upon when scientists assess the consequences of making a mistake while deciding about the evidential support for a hypothesis (Rudner 1953). Furthermore, sometimes scientists have to decide, in a value-laden way, what kind of evidence is relevant for the hypothesis (Longino 1990). Moral and social values can also enter through the background assumptions of scientific reasoning (Longino 1990).

Behavioural scientists, as all scientists, rely on values when it comes to reasoning about possible consequences of making false positive (rejecting a true hypothesis), or false negative (accepting a false hypothesis) errors, while deciding whether evidence is sufficient for supporting or rejecting a hypothesis. (Douglas 2014, p.176). Furthermore, Lacey (2003) points out that in the behavioural and cognitive sciences scientists make choices about adopting, what he calls, behaviourist, cognitivist, or sociobiological strategies. The behaviourist strategy constrains hypotheses to those that concern lawful relations between behaviours and environments, whereas the cognitivist strategy does so to those that concern representations of mental structures and computational accounts of mental processes. Lacey notices that ‘radical behaviorist approaches are partly motivated by the value of furthering our capability to exercise control over human behaviour, and some cognitive psychology approaches are motivated partly by highlighting the values of rationality and freedom’ (Lacey 2003, p.219).

All this means that scientific evidence from the behavioural sciences that enters policy can be already value-laden (even though still reliable and objective). Therefore, apart from testing the effects of nudge interventions advocated by proponents of nudging, it is also important to be able to scrutinize the values that in various ways enter behavioural research. These values should be made subject to debate, especially when the scientific findings are used in the policy context. The next section (Safeguards) suggests how this could be done. Before we move there, let us see which approaches within the behavioural sciences inform contemporary behavioural policies.

Behavioural sciences do not form an unified, integrated body of knowledge. They comprise studies in decision theory, intertemporal choice, social and cognitive psychology, evolutionary approaches, to name only few (see Malecka and Nagatsu 2018 on the variety of approaches within the behavioural sciences). Behavioural policy instruments are informed only by a sub-set of behavioural research: cognitive psychology and behavioural economics (which itself developed out of cognitive psychology, e.g. Tversky and Kahnemann 1975) and use experimental methodology. Gigerenzer (2015) and his critique of the evidence base of nudging is an example for scepticism about the selectivity of the approach; but even this intradisciplinary critique has to be seen in context. With increasing popularity, we hypothesize that the scholarly challenges from other approaches within the behavioural sciences will only increase. It is likely to assume

that more scholars will challenge the underlying scientific foundations of the behaviour change programme from within the behavioural sciences (broadly understood): Already, different schools within psychology have begun to challenge parts of the behaviour change agenda – see the recent debates between evolutionary and social psychology and their institutional representation (Duarte et al. 2015). Arguably, with the sustained popularity of nudge approaches, there will be both increased interdisciplinary scrutiny and intradisciplinary critique.⁹

Other research programmes within the behavioural sciences, that could provide useful evidence for policy-making in the areas where nudges are being proposed (e.g. health) are not included: references to socio-economic research on health behaviour are neglected by proponents of nudges (Quigley 2013), as are prior attempts of psychologists to systematically influence policy (Fawcett et al. 1988). This selective interdisciplinarity does not acknowledge considerable diversity within the contributing disciplines in advocacy of behavioural applications. Favouring of one approach over another can be related to the practical aim and attempt of controlling behaviour, or steering it into more rational directions (Lacey 2003).

5.1 Safeguards

Scientific knowledge that is a basis for behavioural change policies can be value-laden in the sense explained above. Values that sneak into behavioural research can influence the ways in which this research is being applied to policy. Certain approaches in the behavioural sciences become more influential than others. This can be related to the ways in which epistemic and non-epistemic values and interests (e.g. in controlling human behaviour) are entangled. Philosophers of science's response to the question of how to deal with value-ladenness of scientific research have suggested to 'democratize science'. One way to achieve this democratization is to make values embedded in the scientific process explicit and open for debate. The proposal here is that the public receives the chance to legitimately contest (1) the direction of scientific research effort, (2) the legitimacy and acceptability of expertise, (3) the institutional structures for science (public assessment of research agendas, of expertise, of science's institutions) (Douglas 2009).

Citizens need to be able to trust experts, but are often unable to follow scientific reasoning. Anderson (2011) has come up with criteria on how scientific integrity and the trustworthiness of those making scientific claims can be safeguarded. The danger of approaching more and more fields 'through science' without rigorous justification, caution and sophistication (see Pigliucci on scientism and the debate on scientific imperialism generally in Mäki et al. 2017) is that the scientific integrity of experts is exposed.¹⁰

Behavioural change agents can profit from such institutional approaches that ask about the responsibility of groups of behavioural change agents for democracy and politics (see also Thompson 2005). Currently, the way in which behaviour change policies are implemented does not offer institutionalized structures on any of these three aspects. Scrutinizing values embedded in behavioural research can have significant impact on the ways in which this research is applied; it can also allow for more diversity of approaches informing behavioural policies. In addition, the construction of nudge units in a way that requires seeking advice from experts coming from different research traditions within the behavioural sciences should guarantee more pluralism in attempts to find policy solutions.

The requirement of having an expert in the philosophy of science in nudge units could also bring more meta-reflection on the knowledge problems that reliance on the scientific evidence in policy-making provokes. The interaction between philosophers and behaviour change practitioners is surprisingly active already – but an institutional perspective is rarely heard in nudge units.

6. CONCLUSION

While concluding, we want to summarize our tentative suggestions. These aim to stimulate discussion on how an institutional perspective can lead to safeguards. We collect broad ideas for how to translate theoretical misgivings in each section with constructive proposals. They are very preliminary, and should start rather than close debate (see Table 24.1).

The aim of our text is to defend the claim that the behavioural instruments are in tension with law, politics and science. Our work is an example of an institutional perspective on the behaviour change. As such, the analysis is at a high level of abstraction, which is indeed a criticism levied against ours and similar criticisms by Sunstein (2015) who holds that scholars ‘tend to get carried with’ abstract concerns, creating unnecessary confusion unless critique is brought to bear on particular practices. We agree with Sunstein that some critical commentaries can be accused of this. Nevertheless, it is important, if not vital, to think about these abstractions, even in highly generalized ways. This is because whether a theoretical worry will play out in practice depends on the type of behavioural applications actually used in practice. Our framework does not provide criteria for the

Table 24.1 How to safeguard ...

Law	Politics	Science
Regulating nudge units themselves	Expiration dates of behavioural instruments that lead to recurring debates in parliaments	Including different perspectives on human behaviour (e.g. include different schools of psychology in nudge units)
Behavioural disclosure statements to counter non-cognitive effects	Understanding nudge units as political actors (aid from political scientists)	Interdisciplinary evaluation of nudges (multidimensional impact assessments)
Oversight by nudging oversight bodies; nudging ombudsman	Ideological and political diversity in nudge units	Increased meta-reflection by bringing philosophy of science insights
Clarifying limits of administrative nudging	‘Democratizing nudge units’ (e.g. ownership structure, issue selection)	Interdisciplinary review boards
Wider liability for nudge architects	Engaging with critics	Making values explicit
Extensive nudge registries	Making nudges reversible, enable preconditions for resistance	Democratize scientific expertise (public assessment of research agendas, institutions, training)

adjudication of individual interventions. We rather aim at something else. We want to convince the reader that certain institutional features are worth preserving, and are in tension with behaviour change. We believe that these concerns have been overlooked by proponents of behaviour change. The proponents of behaviour change have gotten carried away with the design and implementation of behaviour change without seeing the broader political, institutional and societal context in which the proposal of behaviour change is advocated.

ACKNOWLEDGEMENT

Magdalena Małecka's work was sponsored by Academy of Finland grant no. 308682 and by the Marie Skłodowska-Curie Fellowship EPISTEMEBEHAVIOUR. The authors would like to thank Silke Beck and Holger Strassheim for their invitation to the workshop in Pisa where the first draft of this chapter was presented and discussed. We are also grateful to the workshop's participants for their useful comments.

NOTES

1. In what follows, we draw on Lepenies and Małecka (2015) and Lepenies and Małecka (2016).
2. We here use 'planner' interchangeably with 'policy-maker' and 'choice architect'.
3. While this is true for many other policy instruments, RCTs are heralded as informing social policy in a particularly 'scientifically rigorous' way. To ensure internal validity, the scope for citizens participating in RCTs as active participants rather than subjects to be studied, is necessarily limited (though this is not impossible).
4. We try to abstain from philosophical debates about the normativity of law. We merely say that law which is normative (binding) provides reason for action, and that it approximates one of our ideal types. We do this without defending a theory of normativity that may underlie it. We simply examine what acceptance or elimination of normativity of law leads to for our understanding of agency, as well as for a given legal and political system.
5. Sunstein notes however: 'In many societies, legislatures have explicitly required or authorized nudges, as in the cases of automatic enrollment in savings plans, calorie labels, energy efficiency labels, automatic voter registration, and graphic warning labels for cigarettes' (Sunstein 2016, p. 73).
6. Sunstein (2016) is sceptical about the need to show the exact psychological mechanisms at work in policy. 'Is it manipulative not to be transparent about the psychological mechanisms that make influences work?... If the act is itself transparent, and if deliberative capacities are sufficiently involved, then a failure to tell people about the underlying psychological mechanisms does not mean that manipulation is necessarily involved. For government action, however, reason-giving is ordinarily required, and reason-giving should include an explanation of why a particular form of influence has been chosen – which includes an account of psychological mechanisms' (Sunstein 2016, p. 105).
7. Note also how knowledge of the presence of non-cognitive policy instruments in a given policy field more generally might unsettle citizens: it might undermine their trust in traditional policy-making as they cannot tell if behavioural insights are used (on them) or not.
8. This section draws on arguments made in Małecka and Lepenies (2018).
9. All this is being compounded by the fact that nudge units are active contributors to research, too.
10. There might be some 'behavioural mission creep': autonomous nudge units giving justifications invoking 'science' but instead venturing into new policy fields with new methods without democratic or legal checks and balances. This matters because in practice, those nudge units with a 'dedicated and specialized team, highly informed and specifically devoted to the relevant work, and with expertise in the design of experiments' have been most influential. Sunstein (2014, p. 588): 'Authority greatly matters. The UK has had the most experience with this kind of approach, and its Behavioural Insights Team has succeeded in part because it has enjoyed high-level support and access.'

REFERENCES

- Alemanno, A. and A.L. Sibony (eds) (2015), *Nudge and the Law: A European Perspective*, Oxford: Hart Publishing.
- Alemanno, A. and A. Spina (2014), 'Nudging legally: on the checks and balances of behavioral regulation', *International Journal of Constitutional Law*, **12** (2), 429–456.
- Anderson, E. (2011), 'Democracy, public policy, and lay assessments of scientific testimony', *Episteme*, **8** (2), 144–164.
- Boudry, M. and M. Pigliucci (eds) (2017), *Science Unlimited?: The Challenges of Scientism*, Chicago: University of Chicago Press.
- Cserne, P. (2015), 'Making sense of nudge-scepticism: three challenges to EU law's learning from behavioural sciences', in A. Alemanno and A.L. Sibony (eds), *Nudge and the Law: A European Perspective*, London: Hart Publishing, pp. 279–300.
- Cserne, P. (2016), 'Is nudging really extra-legal?', *The Tocqueville Review/La Revue Tocqueville*, **37** (1), 159–180.
- Douglas, H. (2014), 'Values in social science', in N. Cartwright and E. Montuschi (eds), *Philosophy of Social Science: A New Introduction*, Oxford: Oxford University Press, pp. 162–182.
- Douglas, H. (2009), *Science, Policy, and the Value-Free Ideal*, Pittsburgh, PA: University of Pittsburgh Press.
- Duarte, J., J. Crawford, C. Stern, J. Haidt, J. Lee and P. Tetlock (2015), 'Political diversity will improve social psychological science', *Behavioral and Brain Sciences*, **38**, 1–58.
- Fawcett, S.B., Bernstein, G.S., Czyzewski, M.J., Greene, B.F., Hannah, G.T., Iwata, B.A., Jason, L.A., Mathews, R.M., Morris, E.K., Otis-Wilborn, A. and Seekins, T., 1988. Behavior analysis and public policy. *The Behavior Analyst*, **11** (1), 11–25.
- Freeman, S. 2007. *Rawls*. London: Routledge.
- Fuller, L.L. (1964), *The Morality of Law*, New Haven: Yale University Press.
- Gigerenzer, G. (2015), 'On the supposed evidence for libertarian paternalism', *Review of Philosophy and Psychology*, **6** (3), 361–383.
- Hallsworth, M. (2016), 'Seven ways of applying behavioral science to health policy', in C.T. Robertson, I. Glenn Cohen and H. Fernandez Lynch (eds), *Nudging Health: Health Law and Behavioral Economics*, Baltimore, MD: Johns Hopkins University Press, pp. 40–51.
- Halpern, D. and D. Mason (2015), 'Radical incrementalism', *Evaluation*, **21** (2), 143–149.
- Hansen, P. and A. Jespersen (2013), 'Nudge and the manipulation of choice: a framework for the responsible use of the nudge approach to behaviour change in public policy', *European Journal of Risk Regulation*, **2013** (1), 3–28.
- Hausman, D.M. and B. Welch (2010), 'Debate: to nudge or not to nudge', *Journal of Political Philosophy*, **18** (1), 123–136.
- House of Lords (2011), *Science and Technology Select Committee Behaviour Change HL Paper 170*. Accessed 13 January 2019 at <https://publications.parliament.uk/pa/ld201012/ldselect/ldsctech/179/179.pdf>.
- John, P., G. Smith and G. Stoker (2009), 'Nudge nudge, think think: two strategies for changing civic behaviour', *Political Quarterly*, **80** (3), 361–370.
- Jolls, C., C.R. Sunstein and R. Thaler (1998), 'A behavioral approach to law and economics', *Stanford Law Review*, **50** (5), 1471.
- Jones, R., Pykett, J. and Whitehead, M. 2011. The geographies of soft paternalism in the UK: the rise of the avuncular state and changing behaviour after neoliberalism. *Geography Compass*, **5**(1), pp.50–62.
- Kemmerer, A., C. Möllers, M. Steinbeis and G. Wagner (eds) (2016), *Choice Architecture in Democracies – Exploring the Legitimacy of Nudging*, Oxford/ Baden-Baden: Hart and Nomos.
- Kourany, J.A. (2010), *Philosophy of Science after Feminism*, Oxford: Oxford University Press.
- Lacey, H. (2003), 'The behavioral scientist qua scientist makes value judgments', *Behavior and Philosophy*, **31** (1), 209–223.
- Leggett, W. (2014), 'The politics of behaviour change: nudge, neoliberalism and the state', *Policy and Politics*, **42** (1), 3–19.
- Lepenieš, R. and M. Małecka (2015), 'The institutional consequences of nudging – nudges, politics, and the law', *Review of Philosophy and Psychology*, **6** (3), 427–437.
- Lepenieš, R. and M. Małecka (2016), 'Nudges, Recht und Politik: Institutionelle Implikationen', *Zeitschrift Für Praktische Philosophie*, **3** (1), 487–530.
- Longino, H.E. (1990), *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*, Princeton, NJ: Princeton University Press.
- Lourenço, J.S., E. Ciriolo, S. Rafael Almeida, X. Troussard and European Commission Joint Research Centre. (2016), *Behavioural Insights Applied to Policy: European Report 2016*, Luxembourg: Publications Office.
- Małecka, M. and R. Lepenieš (2018), 'Is the behavioural approach a form of scientific imperialism? An analysis

- of law and policy', in U. Mäki, A. Walsh, and M. Fernández Pinto (eds), *Scientific Imperialism: Exploring the Boundaries of Interdisciplinarity*, Oxford: Routledge, pp. 254–273.
- Malecka, M. and Nagatsu, M. (2018), 'How behavioural research has informed consumer law: the many faces of behavioural research' in K. Purnhagen, A. Sibony and H. Micklitz (eds), *Research Handbook on Methods in Consumer Law*, Cheltenham, UK and Northampton, MA: Edward Elgar Publishing, pp. 357–398.
- Mäki, U., A. Walsh and M.F. Pinto (eds) (2017), *Scientific Imperialism: Exploring the Boundaries of Interdisciplinarity*, Oxford: Routledge.
- McCrudden, C. and J. King (2016), 'The dark side of nudging: the ethics, political economy, and law of libertarian paternalism', in A. Kemmerer, C. Möllers, M. Steinbeis and G. Wagner (eds), *Choice Architecture in Democracies — Exploring the Legitimacy of Nudging*, Oxford/ Baden-Baden: Hart and Nomos, pp. 75–139.
- McMahon, J. (2015), 'Behavioral economics as neoliberalism: producing and governing homo economicus', *Contemporary Political Theory*, **14** (2), 137–158.
- OECD (2017), 'Behavioural insights and public policy: lessons from around the world.' Accessed 2 January 2018 at http://www.oecd-ilibrary.org/governance/behavioural-insights-and-public-policy_9789264270480-en.
- Pigliucci, M. (2015), 'Scientism and pseudoscience: a philosophical commentary', *Journal of Bioethical Inquiry*, **12** (4), 569–575.
- Quigley, M. 2013. Nudging for health: On public policy and designing choice architecture. *Medical Law Review*, **21** (4), 588–621.
- Rudner, R. (1953), 'The scientist qua scientist makes value judgments', *Philosophy of Science*, **20** (1), 1–6.
- Schubert, C. (2017), 'Exploring the (behavioural) political economy of nudging', *Journal of Institutional Economics*, **13** (3), 499–522.
- Shafir, E. (2012), *The Behavioral Foundations of Public Policy*, Princeton, NJ: Princeton University Press.
- Solomon, M. (2001), *Social Empiricism*, Cambridge, MA: MIT Press.
- Straßheim, H. and P. Kettunen (2014), 'When does evidence-based policy turn into policy-based evidence configurations, contexts and mechanisms', *Evidence and Policy*, **10** (2), 259–277.
- Sunstein, C.R. (1997), 'Behavioral analysis of law', *The University of Chicago Law Review*, **64** (4), 1175–1195.
- Sunstein, C.R. (2014), *Why Nudge? The Politics of Libertarian Paternalism*, New Haven: Yale University Press.
- Sunstein, C.R. (2015), 'Nudges, agency, and abstraction: a reply to critics', *Review of Philosophy and Psychology*, **6** (3), 511–529.
- Sunstein, C.R. (2016), *The Ethics of Influence*, Cambridge: Cambridge University Press.
- Thaler, R.H. (2015), *Misbehaving : The Making of Behavioral Economics*, London: W.W. Norton.
- Thaler, R.H. and C.R. Sunstein (2008), *Nudge : Improving Decisions about Health, Wealth, and Happiness*, New Haven: Yale University Press.
- Thompson, D.F. (2005), *Restoring Responsibility: Ethics in Government, Business, and Healthcare*, Cambridge: Cambridge University Press.
- Tversky, A. and D. Kahneman (1975), 'Judgment under uncertainty: heuristics and biases', in *Utility, Probability, and Human Decision Making*, pp. 141–162.
- van Aaken, A. (2016), 'constitutional limits to paternalistic nudging: a proportionality assessment', in A. Kemmerer, C. Möllers, M. Steinbeis, and G. Wagner (eds), *Choice Architecture in Democracies — Exploring the Legitimacy of Nudging*, Oxford/ Baden-Baden: Hart and Nomos, pp. 161–195.
- Wylie, A. and L.H. Nelson (2009), 'Coming to terms with the values of science: insights from feminist science', in H. Kincaid, J. Dupré and A. Wylie (eds), *Value-Free Science? Ideals and Illusions*, Oxford: Oxford University Press, pp. 58–86.

