

[Robert Baldwin](#)

From regulation to behaviour change: giving nudge the third degree

**Article (Accepted version)
(Refereed)**

Original citation: Baldwin, Robert (2014) *From regulation to behaviour change: giving nudge the third degree*. [The Modern Law Review](#), 77 (6). pp. 831-857. ISSN 0026-7961

DOI: [10.1111/1468-2230.12094](https://doi.org/10.1111/1468-2230.12094)

© 2014 The Author

This version available at: <http://eprints.lse.ac.uk/64343/>
Available in LSE Research Online: November 2015

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (<http://eprints.lse.ac.uk>) of the LSE Research Online website.

This document is the author's final accepted version of the journal article. There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

From Regulation to Behaviour Change: Giving Nudge the Third Degree

Robert Baldwin*

Key Words: Regulation; Nudge; Behaviour Change.

Behaviour change strategies such as ‘nudge’ have become hugely popular with administrations on both sides of the Atlantic. Nudging, however, is a practice that raises both conceptual and controversial issues and these have to be addressed in examining the conditions under which nudging can be used effectively and acceptably. Hitherto there has been a lack of conceptual clarity in many debates about nudge but a key to understanding nudge-related issues is to distinguish between three very different degrees of nudge. These three degrees raise different, and identifiable, concerns and it is possible to assess the extent to which these worries can be responded to in positive terms. The compatibility of nudging with other control devices cannot be assumed and, when contemplating nudging, it is essential to be clear and open about the philosophical basis for such action as well as to be aware that different modes of intervention may operate with clashes of logic that threaten not only effectiveness but also the serving of representative and ethical ends.

‘Nudging’ involves structuring the choices that people make so as to lead them towards particular outcomes. Placing fruit next to the supermarket till, for example, gives a nudge towards healthy eating. Such ‘behaviour change’ strategies have become hugely popular with administrations on both sides of the Atlantic. In the UK, David Cameron set up a Behavioural Insights Team (or ‘Nudge Unit’) at the centre of UK government in 2010¹ so as to foster alternatives to traditional regulation and to move towards less-restrictive and lower-cost controls of behaviour. The governmental endorsement of nudging has not, however, been based on clear positions regarding the nature of ‘nudge’ or the role of nudge in the array of state control devices.

This article seeks to address these deficiencies and presents three central arguments. First, that, before nudging is further embraced, it is necessary to think much more clearly about the concept of nudge and to distinguish between the different degrees of nudge that can be applied to targets. Second, that different degrees of nudge can present distinct issues of effectiveness and can give rise to separate concerns of a representative and ethical nature. Third, that a number of important issues are easily overlooked if nudge is seen as just another mode of influence in the toolbox of state controls.

* Professor of Law, London School of Economics and Political Science. The author is grateful for the comments of Vanessa Finch, Veerle Heyvaert, Andrew Lang, Martin Lodge, Henry Rothstein and the anonymous referees.

¹ The Behavioural Insights Team (BIT, or ‘Nudge Unit’) was set up in July 2010 with a remit to find innovative ways of encouraging, enabling and supporting people to make better choices for themselves. Professor Richard Thaler was appointed as an external expert to the BIT - See more at: <http://blogs.cabinetoffice.gov.uk/behavioural-insights-team/author/behavioural-insights-team/#sthash.mxq3kUis.dpuf>
In early 2014 the nudge unit was part-privatised by sale to a charity and to employees. See: <http://www.bbc.co.uk/news/uk-politics-26030205>

The article starts by outlining the rise of ‘nudge’ before reviewing the concept of nudge and distinguishing between three different degrees of nudge. It then explores the limits to the effectiveness of the nudge strategy and the representational and ethical problems posed by the different degrees of nudge. The conditions for effective and acceptable use of nudge are then investigated before the role of nudge in relation to other intervention methods is explored. Emphasis will be placed on the dangers involved in seeing nudge as simply as an extra option in the toolbox of control devices. Finally, conclusions will be offered on the feasibility of constructing a conceptual and ethical basis for making choices regarding this fashionable intervention tool.

THE RISE OF NUDGE

The idea of ‘nudge’ is commonly associated with the book of that title published by Richard Thaler and Cass Sunstein in 2008.² The two authors define a ‘nudge’ as: ‘... any aspect of the choice architecture that alters people’s behaviour in a predictable way without forbidding any options or significantly changing their economic incentives. To count as a nudge, the intervention must be easy and cheap to avoid.’³ In this conception the ‘choice architecture’ is the environment that frames an individual’s choice. Thus: ‘Putting the fruit at eye level counts as a nudge. Banning junk food does not.’⁴

The proponents of nudge build on the well-established insights of cognitive psychology and behavioural economics to contend that control systems need to take on board the bounded rationality of citizens when they make daily decisions.⁵ Individuals are seen as constrained by limitations of information, cognitive capacity and self-control and as tending to rely on a series of heuristics and shorthand methods of using information.⁶ As a result, they often make poor decisions that do not serve their welfare. Nudging holds that citizens’ choices tend to be influenced by a host of framing factors – such as the ways in which options are presented – and it advocates structuring choice architectures so that it is easier for consumers or others (such as regulated firms) to act in ways that are beneficial to them.⁷

At the heart of nudging is a philosophy entitled ‘libertarian paternalism’.⁸ Thus, nudge is said to possess a paternalistic dimension in stimulating choices that are seen

² R. Thaler and C. Sunstein, *Nudge* (New Haven: Yale UP, 2008; revised as Penguin 2009). See also the two authors’ ‘Libertarian Paternalism’ (2003) 93 *American Economic Review* 175-179; ‘Libertarian Paternalism is Not an Oxymoron’ (2003) 70 *University of Chicago Law Review* 1159-1202’.

³ Thaler and Sunstein 2008, n 2 above, 6 (page references in this article are to the revised Penguin edition of 2009).

⁴ *ibid.* p.8.

⁵ On such insights see e.g. A. Tversky and D. Kahneman, ‘Judgement Under Uncertainty: Heuristics and Biases’ (1974) 185 *Science* 1124-1131; H. Simon, *Administrative Behaviour* (Fourth Edition) (New York: The Free Press, 1997).

⁶ Thaler and Sunstein 2008, n 2 above, Chapter 1) note many factors that lead to mistakes in human judgement, such as: optimism and over-confidence; loss aversion; a status quo bias; framing; inertia; inattention and error; anchoring; availability; and representativeness.

⁷ In their first example of a nudge, Thaler and Sunstein posit that the food in a school cafeteria is arranged so that students are likely to make the most healthy choices Thaler and Sunstein 2008, n 2 above, 5.

⁸ Thaler and Sunstein 2008, n 2 above, 5-6.

as welfare enhancing for the individual but it combines this with a libertarian aspect in so far as it purports to leave the target person or firm free to choose to take the non-sensible course of action.

Leaving aside incentive regimes, Thaler and Sunstein give examples of at least seven nudge tools.⁹ **Defaults** encourage the healthy choice by making this the given selection unless active steps are taken to avoid this outcome – steps that will not be encouraged by the natural tendency to human inertia. A prime example of a default is a presumption that all citizens consent to be organ donors unless they register their unwillingness to donate (which, as noted, Thaler and Sunstein state, they should be able to do easily). Similarly, employees can be enrolled automatically onto a pension scheme unless they actively opt out. **Persuasive, campaigning and counselling** strategies can influence and shape decision-making. **Design** approaches can be used – as where the main office doors lead directly to the stairs, not the lifts; or the departmental smoking zone is placed at a distance from the work area. **Commitments** to healthy activities, such as eating less or charitable giving can be stimulated so as to encourage delivery on good intentions. **Transactional shortcuts** can be provided so that, for example, a special credit card makes charitable donations easy to make and record. Exemptions, similarly, can be used to spare people from procedures if they do the sensible thing – as where motorcycle helmet users do not have to obtain a licence. **Information mechanisms** can be deployed to inform people of pitfalls or errors, or to present options in ways that foster healthy decisions, or to facilitate corrective actions. Information can also be given on social norms or other people’s performance (e.g. the percentage of citizens who have already completed a tax return)¹⁰. Information can also be mapped and structured (especially in relation to complex issues) so as to provide citizens with easier routes to the sensible decision. **Warnings and reminders** can be given so as to discourage unwise actions such as the sending of intemperate or uncivil emails.¹¹

As noted, the Cameron administration set up the Behavioural Insights Team (the ‘Nudge Unit’) as part of a movement away from command-based regulation and fiscal controls in favour of ‘smarter’ strategies that would use behavioural insights rather than rules or financial incentives to influence behaviour. This drift was seen as continuing the deregulatory thrust of the Better Regulation Executive and as an espousing of ‘non-regulatory means of achieving behaviour change’.¹² As for the merits of a ‘non-regulatory approach’, the Minister for Government Policy, Oliver Letwin MP, stressed four factors: effectiveness, cost-effectiveness, less rigid impositions on individuals and reduced burdens on business.

Since its inception the BIT has investigated the potential of nudges to improve practices relating to such matters as: organ donor registrations, charitable donations, energy saving and smoking.¹³

⁹ Incentive strategies with economic effects are not nudges as defined by Thaler and Sunstein 2008, n 2 above, 8. For a dozen examples of nudges see Thaler and Sunstein (2008) Chapter 14.

¹⁰ See e.g. R. Cialdini, ‘Activating and Aligning Social Norms for Persuasive Impact’ (2006) *Social Influence* 3-15.

¹¹ The example provided at Thaler and Sunstein 2008, n 2 above, 233-234.

¹² House of Lords Science and Technology Select Committee Second Report of Session 2010-12 *Behaviour Change* (HL Paper 179, 2011) (hereafter ‘House of Lords Select Committee, 2011’) para. 5.3 quoting the evidence of the head of the BIT, Dr David Halpern.

¹³ For details of the BIT’s work see its website:

<https://www.gov.uk/government/organisations/behavioural-insights-team>

THE CONCEPT OF NUDGE

Understanding what nudging involves is a pre-condition for assessing the effectiveness and acceptability of nudge as an intervention strategy – or for evaluating any potential use of nudge in a specific context. If the concept of ‘nudge’ is unclear or encompasses a number of very different devices this complicates any attempt to assess its potential performance.

A problem here is that Thaler and Sunstein adopt a conception of the nudge that is both broad and contested by some commentators. Thus, the two authors include within nudging a number of examples of simple information provision (e.g. requirements that cigarette packs display information on the risks of smoking; or demands that firms inform employees of work hazards; or suggestions that red warning lights be used to inform of needs to change air filters¹⁴). They also cloud matters by citing, as examples of nudges, some uses of economic incentives (e.g. to limit pregnancies¹⁵) which contradicts their definitional exclusion of fiscal measures from the category of nudges.¹⁶

On the matter of simple information supply, some commentators would depart from Thaler and Sunstein and omit these activities from the category of ‘nudges’.¹⁷ Thus Peter John and colleagues have (like others) distinguished between ‘nudges’ and actions that are taken so as to foster deliberative decision-making processes.¹⁸ ‘Nudges’, thus conceived, operate by working within the domain of ‘automatic’ (or ‘System 1’) responses since the essence of nudging is to control individuals by exploiting their cognitive and emotional limitations rather than by seeking to improve their capacities to exercise informed, rational and conscious choices. In contrast, interventions that are based on a ‘think’ strategy are designed to encourage ‘reflective’ (or ‘System 2’) decision procedures.¹⁹ ‘Think’ strategies will enhance rationality by supplying information that will be used in reflective decision-making (‘These cigarettes may harm your health’). ‘Simple’ information supply, from this perspective, is thus inconsistent with nudging as it stimulates conscious decision-making.

Those who disagree with Thaler and Sunstein on this point reinforce their position by arguing that simple information provision is not, in any event, covered by the idea of ‘libertarian paternalism’ because actions that are taken so as to enhance an individual’s exercise of ‘reflective’, rational choice, and which respect the decision-

¹⁴ See Thaler and Sunstein 2008, n 2 above, 200–201; 232–233 and D. Hausman and B. Welch, ‘Debate: To Nudge or Not to Nudge’ (2010) 18 *The Journal of Political Philosophy* 123–136, 127.

¹⁵ See Thaler and Sunstein 2008, n 2 above, 232.

¹⁶ See: British Medical Journal, Editorial ‘One Nudge Forward, Two Steps Back’ (2011) 342 *BMJ* d401.

¹⁷ P. John, G. Smith and G. Stoker, ‘Nudge Nudge, Think Think: Two Strategies for Changing Civic Behaviour’ (2009) 80 *The Political Quarterly* 361–370. The House of Lords Select Committee, 2011, similarly took the view, contra Thaler and Sunstein, that: ‘Nudges prompt choices without getting people to consider their options consciously, and therefore do not include openly persuasive interventions such as media campaigns and the straightforward provision of information.’ (para. 2.9). See also L. Bovens, ‘The Ethics of *Nudge*’ in T. Grune-Yanoff and S. Hansson (eds.) *Preference Change* (Berlin: Springer, 2008).

¹⁸ John, Smith and Stoker *ibid*.

¹⁹ See Thaler and Sunstein 2008, n 2 above, 21. More generally see D. Kahneman, *Thinking, Fast and Slow* (London: Allen Lane, 2011).

making autonomy of the individual, are not paternalistic at all – there is no substitution of the judgement of the nudger for that of the individual who is nudged.²⁰ Thaler and Sunstein have, nevertheless, been followed by many policy-makers in including ‘simple information’ and other ‘think’ devices alongside ‘System 1’ influences within the category of nudges.²¹

THREE DEGREES OF NUDGE

The prevalence of the latter, highly inclusive, notion of nudge means that distinctions have to be drawn between different degrees of nudge if there is to be conceptual clarity on nudging. Nudges, it is argued here, vary in the degree to which they impact on the autonomy of individuals as decision-makers and it is useful to distinguish between three levels of such impact. This is because distinct sets of representative, ethical and practical issues are raised by these three degrees of nudge.

It should be noted that the degree of a nudge is a separate matter from the particular nudging tool that is deployed (be that a default rule, an exercise in information supply or some other approach). Thus, a tool such as supplying information may be used, as will be demonstrated, for first, second, or third degree nudging.

In the framework proposed here, ‘First Degree nudges’ respect the decision-making autonomy of the individual and enhance reflective decision-making. Typical First Degree nudges involve the supply of simple information to individuals or the imparting of reminders (‘There are three weeks left to complete the tax return.’). Such nudges can be distinguished from two other, more serious, kinds of intervention.

A ‘Second Degree nudge’ typically builds on behavioural or volitional limitations so as to bias a decision in the desired direction. Thus, a default rule with an opt-out can be used to shape decisions (as in presumed consent to organ donation), or the physical environment for decision-making may be designed to apply influence (as where the office smoking zone is placed at a distance from the work area). Both of these examples nudge individuals’ decisions by relying on human inertia.

The Second Degree nudge involves a greater impact on individual autonomy than the First Degree nudge since the targeted individual’s behavioural or volitional limitations and ‘automatic’ responses will in practice lead him or her to accept the nudge with limited awareness and reflection. It is, nevertheless, the case both that such a nudge will shape an individual’s decision and that the target of the nudge would be capable, *on reflection*, of realising that a nudge has been administered and assessing its broad effect.

A ‘Third Degree nudge’ offers a yet more serious intrusion on autonomy because it involves behavioural manipulation to an extent that other nudges do not. Thus, framing devices can be used to shape the decisions and preferences of an individual in a manner that is resistant to unpacking in so far as assessing the nature and extent of the nudge is not readily achieved by reflection.²² Thus, a campaign to promote healthy eating might focus on certain negative consequences that are associated with

²⁰ See Hausman and Welch, n 14 above, 127.

²¹ See e.g. Behavioural Insights Team, *Behaviour Change and Energy Use* (London: Cabinet Office, 2011) Chapter 2.

²² On preferences being sensitive to the way a choice problem is described or ‘framed’ see: P. Slovic, ‘The Construction of Preference’ (1995) 50(5) *American Psychologist* 364 - 371; A. Tversky and D. Kahneman, ‘The Framing of Decisions and the Psychology of Choice’ (1981) 211 *Science* 453-458; J. Payne, J. Bettman and E. Johnson, ‘Behavioural Decision Research: A Constructive Processing Perspective’ (1992) 43 *Annual Review of Psychology* 87 – 131.

poor diets so as to structure decisions by giving dominance to particular considerations ('Don't lose your looks, junk the junk food!'). Reflection makes the target aware of the emphasis on 'looks' in the example given but not of the extent of preference adjustment involved in expressing the eating decision as one involving a loss, rather than gain; in giving 'looks' a dominant position in decisions about healthy eating; and in applying the emotive power associated with a loss of looks. The result, in cases of preference shaping, is that even reflection makes little difference to the individual – it only reinforces the pursuit of the 'shaped preferences'. In such cases, the Third Degree nudge does not so much push the tiller as rig the compass bearing.

Another form of Third degree nudge will often rely on salience and affect to produce its result. Typically, there is use of the behavioural economist or psychologist's finding that people are influenced by novel, personally relevant or vivid examples of, say, a harm.²³ In such instances the message receiver will be motivated emotionally in a manner that stands in the way of reflection. Thus, in the smoking context, the cigarette pack would show a graphic display of a corpse. (Alternatively, a smoker could be shown a video of a loved one having a heart attack.) The element of 'manipulation' stems from the use of a level of emotional power so that this: 'Blocks the consideration of all options and threatens the agent's ability to act in accordance with her or his own preferences (as opposed to someone else's).'²⁴

The difference between a Second and a Third Degree nudge is that, with the former, the message receiver has the practical potential to uncover the nudge, and assess its extent, by the exercise of reflection whereas, in the latter, there is a material or complete²⁵ blocking of such reflection, or the neutralising of reflection. The distinction between Second and Third degree nudges does not turn on a sharp differentiation of cognitive, volitional and emotional functions and manipulations, nor is it to presuppose a lack of linkage between these functions – the literature of social psychologists and sociologists of emotion offers an extensive discussion of the extent and nature of such links²⁶ – it is suggested here that Third Degree nudges are distinctive in so far as decision framing through emotional or any other form of influence (or combination of influences) is sufficient materially to obstruct reflection or the assessing, on reflection, of the extent, nature and degree of the nudge.²⁷

²³ J. Blumenthal-Barby and H. Burroughs, 'Seeking Better Health Care outcomes: The Ethics of Using the "Nudge"' (2012) 12(2) *Am. J. of Bioethics* 1-10, 4.

²⁴ *ibid.* 5.

²⁵ An example of complete blocking would be the use of subliminal messaging – which imposes the nudge without the nudged individual's awareness at the time or even afterwards. This would be unannounced, unpublicised subliminal messaging - though a government might openly declare that it has a policy of using subliminal messaging to control e.g. a health issue. Thaler and Sunstein argue that they endorse Rawls' publicity principle as a limitation of nudging and that subliminal messaging is, accordingly, not acceptable because it lacks transparency, see: Thaler and Sunstein 2008, n 2 above, 243-246 and the discussion below.

²⁶ See e.g. J. Storbeck and G. Clore, 'On the interdependence of cognition and emotion' (2007) 21 *Cognition and Emotion* 1212-1237; C. Izard, J. Kagan, and R. Zajonc (eds) *Emotions, Cognition, and Behavior* (Cambridge: CUP, 1984); T. Kemper (ed) *Research Agendas in the Sociology of Emotions* (1990); T. Dalgleish and M. Power (eds), *Handbook of Cognition and Emotion* (Chichester: Wiley, 1999); T. Dalgleish and M. Power, *Cognition and Emotion* (Hove: Psychology Press, 2008).

²⁷ The difference between Second and Third degree nudges may thus be a difference of degree (no pun) rather than kind. Second degree nudges may have an emotional dimension and they may also structure choices - the issue is whether the nudge materially impedes reflection or an awareness, through reflection, of the extent, nature and degree of the nudge.

Table 1: Three Degrees of Nudge

	Typical Characteristics	Example	Impact on autonomy
First Degree Nudge	Supply of simple information or a reminder with the aim of improving the target's capacity to make an informed, rational and conscious choice.	<ol style="list-style-type: none"> 1. Health warning on cigarette pack. 2. Reminder to fill in tax return. 	Respects the autonomy of decision-maker and enhances target's rationality.
Second Degree Nudge	Behavioural or volitional limitations are exploited so as to bias decisions in a favoured direction.	<ol style="list-style-type: none"> 1. An opt-out organ donor regime is instituted. 2. The office smoking zone is placed at a distance from the work area. 	The target could, on reflection, unearth the nature and effect of the nudge – but is unlikely to do so because of behavioural limitations and their tendency to exhibit an 'automatic' response.
Third Degree Nudge	Framing strategies, emotional responses or covert techniques are used to influence decisions or shape preferences.	<ol style="list-style-type: none"> 1. A campaign promotes healthy eating with the slogan: 'Don't lose your looks, junk the junk food!' 2. Shocking images are used to control behaviour – as when photographs of lung cancer victims are used to control smoking. 3. Unpublicised subliminal TV messages are used to encourage e.g. healthy eating or abstention from smoking. 	The target is influenced but reflection is obstructed or reflection materially fails to unpack the nature and extent of the decision or preference shaping.

The distinctions between these three degrees of nudge are important in assessing both the potential effectiveness and the acceptability of nudging by different methods in different contexts. (This is the case in spite of our ability to conjure up instances at the borderlines of these categories.) Questions of acceptability will be considered below but, at this stage, it suffices to note that some degrees of nudge present much more serious representational and ethical issues than others. It follows that there are real dangers in adopting a loose conception of nudge. The most serious of these is that the most benignly uncontentious examples of nudge may be highlighted in debates and that this impedes critical reflection on the more controversial aspects and versions of the device.

LIMITS TO THE EFFECTIVENESS OF NUDGE

Nudges are seen, by many politicians, to be an alternative to ‘command’ regulation²⁸ and it is appropriate, accordingly, to measure the nudge against the yardsticks commonly used in assessing regulatory strategies.²⁹ Relevant issues are whether nudge produces properly targeted outcomes effectively and whether nudging serves representative and ethical values acceptably.

On the question of effectiveness, the nudge has practical limitations – some of which are general and some of which apply especially to certain degrees of nudge. The first limitation to be noted is one that applies to all degrees of nudge and stems from nudge’s strategy of aiming to achieve results by focussing on the decision-making of the individual. It has been pointed out that such a focus may fail to address the causes of, or provide the solution to, a number of problems.³⁰ Obesity, for example, may be seen by nudgers as the result of an accumulation of poor individual decisions (which can be improved by nudges) but this vision fails to take on board a host of biological, social and cultural causes of obesity. It has been argued, for instance, that the rise in obesity of recent years is not to be explained in terms of personal responsibility but in: ‘...a toxic combination of readily available cheap high energy food and drink, fewer opportunities for manual labour, an increase in car ownership, changing social norms concerning cooking and eating, and other features of the “obesogenic” environment.’³¹ If personal consumer decisions are only one element in the causation of obesity, nudges of those decisions are likely to prove of limited effect in dealing with the essential problem. It is likely that obesity has to be

²⁸ See the evidence of the Minister for Government Policy, Oliver Letwin, to the House of Lords Select Committee, 2011, Q. 703, reported at para. 5.3.

²⁹ On yardsticks for measuring regulation see generally R. Baldwin, M. Cave and M. Lodge, *Understanding Regulation* (Second Edition) (Oxford: OUP, 2012) Chapter 3.

³⁰ See: J-F Menard, ‘A “Nudge” for Public Health Ethics: Libertarian Paternalism as a Framework for Ethical Analysis of Public Health Interventions?’ (2010) 3 *Public Health Ethics* 229-238.

³¹ Evidence of Professor Thomas Baldwin, House of Lords Select Committee, 2011, para. 7.4. Other witnesses echoed this point: see the evidence of Dr Ian Campbell, Medical Director, Weight Concern (noted at para.7.8). See also: Foresight and BIS, *Tackling Obesity: Future Choices* (2007). For a discussion of behaviour change interventions and obesity see Appendix 4 of House of Lords Select Committee, 2011. For similar arguments in relation to reducing car use see paras. 7.26 –7.48 and Appendix 5.

addressed with a combination of personal, normative, cultural and environmental measures (regulatory and non-regulatory) rather than left to nudging alone.³²

Nudges that are aimed at individuals, moreover, will not always prove effective when the undesirable behaviour at issue is the product of collective processes and policies (for example where competing interests and pressures within an enterprise produce a corporate decision that harms the environment).³³ Individually - targeted nudges are likely to impact on collectively-generated decisions only when the latter are sensitive to changes in the targeted individuals' decisions (as where consumers are given information on corporate environmental records, this affects sales of the corporation's products, and the corporation is sufficiently responsive to amend its environmental strategies as a result). It follows that, if many modern-day problems are organizationally generated, governmental inclinations to avoid the 'nanny state' by relying on 'non-regulatory' controls and individual responsibilities may hit the mark only under certain conditions.

A second limitation of effectiveness can be expected to vary across degrees of nudge. It is that individuals' responses to nudges will differ across divergent institutional, social, economic and cultural contexts.³⁴ A nudge to encourage recycling may not work as well in a disadvantaged community as in an affluent district.³⁵ Nor will automatic enrolment into a retirement plan (with opt-out) operate to increase savings in the body of the population that does not earn enough to be able to put this money aside.³⁶

Nudges, and different degrees of nudge, accordingly, can be expected to echo the tendency of traditional regulatory styles to vary in their effectiveness when targeted at different types of regulated concern and at people with different characteristics. Table 2 below looks at four examples of nudges as applied to four different models of individual. It indicates whether the effect of a given nudge on the particular target is likely to be high, medium or low. The objective of the discussion here is not to give an exhaustive (and exhausting) account of variations in effectiveness across the huge number of possible ways to implement different nudge tools. The aim is to illustrate the kinds of variation in effectiveness that can be expected across not only degrees of nudge but across different nudge tools (even tools of the same degree). The four nudges set out in Table 2 are thus chosen to exemplify different degrees of nudge but also different nudge tools of the same degree.

The breakdown of individuals used in Table 2 is one familiar in the regulation literature³⁷ and refers to those characterised respectively as: well-intentioned and high capacity; well-intentioned and low capacity; ill-intentioned and high capacity; ill-

³² House of Lords Select Committee, 2011, paras. 7.6 and 8.14 and evidence of Dr M. Hillsdon.

³³ In contrast are instances in which the relevant decisions are made by individuals and those decisions are nudged. In such cases the nudging of individuals can produce reactions on the part of corporations – as where citizens are nudged to smoke fewer cigarettes and corporate producers adapt their marketing strategies as a result.

³⁴ See: e.g. L. Bovens, 'Nudges and Cultural Variance' (2010) 23 *Know.Tech. Pol.* 483-486.

³⁵ John, Smith and Stoker, n 14 above, 369.

³⁶ Menard, n 30 above, 235.

³⁷ See e.g. J. Black and R. Baldwin, 'When Risk-Based Regulation Aims Low: A Strategic Framework' (2012) 6 *Regulation and Governance* 131-148; R. Baldwin, M. Cave and M. Lodge, *Understanding Regulation 2nd ed.* (Oxford, 2012) pp. 230-231; R. Kagan and J. Scholz, 'The Criminology of the Corporation' in K. Hawkins and J. Thomas (eds), *Enforcing Regulation* (Boston: Kluwer-Nijhoff, 1984).

intentioned and low capacity. ‘Intentioned’ here refers to the extent to which the targeted person espouses the same objective as the nudger. Thus, a well-intentioned target will be interested in, and strongly disposed in favour of the behaviour that the nudger sees as virtuous. Ill-intentioned targets will not espouse such objectives and may even be actively opposed to behaving in the manner that the nudger sees as virtuous.

‘Capacity’ refers to the ability of that person to gain, receive absorb and act on information. A high capacity individual will thus be able to receive messages (even complex ones) effectively and will be able to adjust his or her behaviour in response to such messages. Whether such an individual will be *inclined* to adjust their behaviour in accordance with the message received will, of course, depend on their intention rather than capacity. High capacity individuals who are not well-intentioned and not inclined to act in accordance with a message will be very able to adjust their behaviour so as to reject that message and act in ways that are inconsistent with it. Such individuals will possess a high ability to ‘unearth’ nudges, such as defaults, and to resist these. Low capacity individuals will struggle to absorb and act on even simple messages, even when disposed so to act. Low capacity individuals who are ill-intentioned, will, moreover, have very limited ability to adjust their behaviour so as to reject messages that they disagree with and to act in ways that are inconsistent with such messages. They will, in turn, possess poor abilities to ‘unearth’ nudges such as defaults, and resist these.

Table 2: The Anticipated Impact of Nudges on Different Targets.

	<u>Information Mechanism:</u> Simple nutritional information is provided on food packaging. (A First Degree Nudge)	<u>Default:</u> An opt-out regime is adopted for organ donation. (A Second Degree Nudge)	<u>Design:</u> The smoking zone is placed at a distance from the office work area. (A Second Degree Nudge)	<u>Warning:</u> Shocking photographs of lung cancer victims are used to control smoking. (A Third Degree Nudge)
Well-intentioned and High Capacity	High	High	High	High
Well-intentioned and Low Capacity	Low	High	High	High
Ill-intentioned and Low Capacity	Low	High	Medium	High
Ill-intentioned and High Capacity	Low	Low	Low	Medium

Empirical evidence on the above indications is, as yet, thin on the ground (a matter to be returned to below) but, if there is some truth in the estimations summarised in Table 2, this offers at least three relevant messages for potential nudgers. First, different degrees of nudge, different nudge tools, and different ways of implementing given tools can be expected to bring different effectiveness profiles. Thus, the successful use of a First Degree nudge, such as the provision of simple information, can be expected to be far more dependent on the capacity of the target than a Second Degree nudge such as a design, or a default approach or a Third Degree nudge such as a shock image. The First Degree nudge, after all, depends on stimulating high quality conscious decision-making. With targets of low capacity it can be expected that Second and Third Degree nudges will bring special potential – though it can also be seen in Table 2 that even across different tools of the same degree there may be variations in effectiveness with regard to low capacity targets.

A second message is that even a First Degree nudge, such as simple information supply, may prove effective only with a small minority of targets. Third, it can be anticipated that targets who are ill-intentioned and high capacity will tend to be difficult to nudge with a high degree of effectiveness whatever degree or tool of nudge is employed.³⁸

At this point it should be emphasised that the above discussion does not constitute an argument against nudging *per se* but it offers a caution that the respective limitations of different degrees, tools and modes of implementing nudges need to be borne in mind if the nudge approach is to be deployed astutely.

A further limitation on the effectiveness of nudging is presented by the possibility of the counter-nudge, and here there is a contrast with traditional command approaches. If legislators ban the sale of alcohol above a certain strength, they deal with the issue definitively. If they rely on a nudge to control this activity (e.g. imposing a requirement that a person registers before purchasing such alcohol) there is the potential for interested parties to ‘counter-nudge’.³⁹ Thus, the supermarkets have the chance to respond by stacking high strength alcohol next to the till or by advertising it more intensively. The marketing and advertising sectors of the private sector may, indeed, prove formidable opponents for the governmental players of the nudge game. It is unlikely, for instance, that nudgers who seek to control obesity will ever match the power to influence that is provided by the promotional budgets of the fast food industry. This is a point that applies across the range of nudge categories, though the higher the degree of nudge used, the more severe will be the response needed to contest this.

When nudging operates over time, there is another issue that goes to the effectiveness of the device. Nudges that are genuinely paternalistic (Second and Third Degree nudges) may undermine relationships that are based on trust when the practice

³⁸ Table 2 suggests that, even with the default and design strategies that are discussed, targets of ill-intention and high capacity may use their cognitive resources to unmask the nudges and circumvent these. See also the discussion of the limits of regulating internet users’ behaviour through ‘code’ in A. Murray, ‘Internet Regulation’ in D. Levi-Faur (ed), *Handbook on the Politics of Regulation* (Cheltenham: Edward Elgar, 2011).

³⁹ It is conceded that the targets of legislation may respond by strategies such as creative compliance or breaching the law – what they cannot do is counter-legislate.

of nudging is uncovered. Thus, in the field of medicine it has been warned that nudges can be seen as scare tactics by patients and that this can alienate them from their doctors. The use of information selectively or emotively can thus act to the detriment of treatment in the longer term.⁴⁰ Similarly, it has been argued that nudges may lead patients to see their physicians as operating a ‘policing’ function and that this may lead the former to withhold important medical information from those who treat them.⁴¹ The problem will be all the greater if the nudger is seen to be manipulating the truth in ‘constructing a narrative’ for the purposes of nudging – as, for example, through an inaccurate or ‘aspirational’ representation of general practice such as: ‘Nearly all patients consent to this procedure.’

As for nudge’s effectiveness as a device for avoiding need for legal rules, this is easy to exaggerate. Governments can nudge individuals directly but they will often use businesses as their nudging agents for all degrees of nudge. This occurs, for example, when retailers are legally required to provide consumers with certain messages or items of information. In such regimes it is true that consumers do escape from the clutches of command and control laws but this is by no means the case for the affected businesses.

As far as producing *properly targeted outcomes* effectively, it cannot be assumed that nudges (of whatever degree) are any less prone to the triggering of unintended consequences than other forms of regulatory or non-regulatory intervention.⁴² Thus, it has been pointed out that consumers may override intuitive responses or react perversely to nudges.⁴³ Food that is labelled as healthy may lead to excessive consumption.⁴⁴ Positive information on low levels of a certain unhealthy ingredient may blind the consumer to high quantities of another unhealthy component,⁴⁵ and a positive message about one aspect of a food may induce underestimations of other, harmful, qualities.⁴⁶

A further targeting challenge arises when opinions differ or develop on what such outcomes involve – and these challenges may differ across degrees of nudge. First Degree nudges encourage conscious deliberation on the ‘think’ model and operate with a mechanism that can cope with differences and changes in perceptions regarding desired outcomes (e.g. new ideas about the right balance between alcohol and carbohydrate consumption). The danger with nudges that influence automatic responses (Second and Third Degree nudges) is that it is the paternalistic nudger that defines desired outcomes and this may lead to a false sense of security concerning the desirability of those outcomes.⁴⁷ The problem is that, even if the nudger’s views on

⁴⁰ Blumenthal-Barby and H. Burroughs, n 22 above, 5.

⁴¹ *ibid.* 2.

⁴² An assumption apparently made by the Minister, Oliver Letwin MP, in giving evidence – see House of Lords Select Committee, 2011, para. 5.7.

⁴³ See O. Amir and O. Lobel, ‘Stumble, Predict, Nudge’ (2008) 108 Col. L. Rev. 2098-2137, 2116. Ill-intentioned, high capacity targets may be especially prone to perverse reaction.

⁴⁴ House of Lords Select Committee, 2011, para. 7.11.

⁴⁵ An investigation by *The Daily Telegraph* in March 2014 revealed that foods promoted as ‘low fat’ and ‘low calorie’ often had more harmful effects on health than their full fat equivalents. The campaign group Action on Sugar protested that manufacturers were ‘misleading’ shoppers. See <http://www.telegraph.co.uk/health/healthnews/10668189/Low-fat-foods-stuffed-with-harmful-levels-of-sugar.html> (Last visited 8.3.14).

⁴⁶ T. Marteau *et al* ‘Judging Nudging: Can Nudging Improve Population Health?’ (2011) 342 *BMJ* 263, 264.

⁴⁷ John, Smith and Stoker, (2009), n 14 above, 368.

outcomes coincided initially with those of the consumers or citizens being protected, the two strands of thought may, over time, diverge because such nudges are based on paternalism rather than the fostering of an ongoing dialogue about objectives. In such cases, nudge may produce results other than those that the nudged parties would have sought.

Finally, it should be noted that numbers of commentators have stressed that there has emerged very little concrete evidence on the effectiveness of nudge strategies. To an extent this may stem from the paucity of studies that deal with the performance of nudges.⁴⁸ The House of Lords Science and Technology Committee, indeed, stated in 2011 that it had been provided with no examples of significant change in the behaviour of a population having been achieved by non-regulatory measures alone.⁴⁹ Its report concluded that non-regulatory measures were often not likely to be effective when used in isolation and that, usually, a range of policy tools, of both kinds, would be needed to bring about change effectively.⁵⁰ The Committee stressed that more rigorous steps should be taken within government to ensure the effective evaluation of behaviour change interventions.⁵¹

REPRESENTATIONAL AND ETHICAL ISSUES

Nudge, with its philosophy of ‘libertarian paternalism,’ gives rise to a number of representational and ethical concerns. A first of these is that many decisions to nudge lack the transparency and public consideration that are normally associated with command regimes.⁵² If a government issues a law that prohibits citizens from smoking in public places, this is a mode of control that is open, discussed, and implemented after representative processes have been followed.⁵³ If nudging is used, the process used to effect a nudge may be far more secretive—the nudge, for instance, may flow from an administrator’s decision on the design of a public building: a decision not subjected to advanced disclosure or debate. As Luc Bovens has said, nudges ‘typically work better in the dark’.⁵⁴ This may be seen as relatively unobjectionable in the case of First Degree nudges, since these contribute positively

⁴⁸ Marteau, n 26 above, 264.

⁴⁹ House of Lords Select Committee, 2011, para. 5.6.

⁵⁰ *ibid.* paras. 5.13; 8.14.

⁵¹ *ibid.* Chapter 6 and paras. 8.21-8.22. Others have suggested that, given the variety of nudges, such work should deal with a series of issues: What works? For whom? In what circumstances? For how long? How effective would nudges be compared to other interventions? See e.g. Marteau and colleagues at Marteau (2011) n 26 above, 264.

⁵² See R. Brownsword, ‘Code, Control, and Choice: Why East is East and West is West’ (2005) 25 *Legal Studies* 1-21 who contrasts the transparency of ‘command’ (‘East coast’) and ‘design’ (‘West coast’) controls. See similarly: L. Tien, ‘Architectural Regulation and the Evolution of Social Norms’ (2004) 9 *International Journal of Communications Law and Policy* 1. On control through ‘design’, ‘code’ or architecture’ more generally see L. Lessig, *Code and Other Laws of Cyberspace* (New York: Basic Books, 1999); A. Murray and C. Scott, ‘Controlling the New Media: Hybrid Responses to New Forms of Power’ (2002) 65 *MLR* 491.

⁵³ The transparency of ‘command’ approaches should not, however, be exaggerated – many commands are promulgated through delegated legislation and other varieties of soft law. See generally R. Baldwin *Rules and Government* (Oxford: OUP, 1995). It is arguable, nevertheless, that nudges can be instituted by methods that are far less transparent than even the processes used to introduce soft-laws.

⁵⁴ See Bovens, n 14 above, 3 (referring to nudges that would be Second Degree nudges in the framework proposed by this article).

to conscious individual decision-making, but a lack of transparency regarding Second and Third Degree nudges would be far more contentious.

It is true that, where legislation is used to institute the nudge – as when a new law demands that supermarkets place fruit next to the till – there will be the normal public debate that attaches to the legislative process. Many nudges, however, will not demand legislation and may be triggered administratively, as may, for instance: informational nudges, administrative default rules, publicity campaigns, alterations in the physical environment and the messaging of social norms.

The proponents of nudge might protest that nudges of all degrees can be disclosed. Thaler and Sunstein, indeed, state that they object to nudging by subliminal messaging because they endorse Rawls' publicity principle as a limitation of nudging and subliminal messaging is, accordingly, not acceptable because it lacks transparency.⁵⁵ The problem, however, is that, although it is possible for a government to state openly that it is using Second or Third Degree nudges in an area, there are reasons to be pessimistic about the prospects of this occurring so that this enhances rational decision-making or transparency. First, there is the Bovens point, noted above – if the students are told that the cafeteria food is arranged in order to push them towards a healthy diet, this is likely to undermine the effectiveness of the nudge.⁵⁶ Second, a general disclosure about nudging ('We have designed this working environment to help you make healthy decisions') may prove unhelpful in rendering the nudge's existence or nature evident at the time when a particular decision is taken.⁵⁷ After all, the critics would say, nudges (of the Second Degree and Third Degree) are applied on occasions that allow the exploitation of an individual's weaknesses of willpower, emotion and rationality and this is not a context likely to encourage advertence to the fact or extent of a nudge. Showing the drowning bather a flipchart about learning to swim may be poor timing.

In further defence of the nudge it might be suggested that there is little need for debate when all the nudger is doing is making targets' lives 'longer, healthier and better' and trying to 'influence choices in a way that will make choosers better off *as judged by themselves*'.⁵⁸ The nudge, on this view, conduces to the decision that the target would have arrived at had he or she paid full attention and possessed complete information, unlimited cognitive abilities, and complete self-control.⁵⁹ The limitation of this argument is that Second and Third Degree nudges do not involve seeking to identify the conception of welfare that the target would espouse.⁶⁰ In contrast, when nudging goes beyond the supply of simple information, or reminders, at the 'First Degree' level (which can be said to enhance the target's conscious decision-making) what is involved is the nudger acting in a semi-covert or covert manner to further the nudger's own conception of the target's welfare – which may involve re-shaping the target's idea of their own welfare.

⁵⁵ Thaler and Sunstein 2008, n 2 above, 243-246.

⁵⁶ See Bovens, n 14 above, 3-4.

⁵⁷ Brownsword, n 52 above, 14 - 17 makes the additional point that 'design' approaches can 'embed' solutions so that those affected in the future have little consciousness of prior design decisions and no input into the ways in which they are controlled.

⁵⁸ See Thaler and Sunstein 2008, n 2 above, 5.

⁵⁹ *ibid.* 6.

⁶⁰ On the argument that targets' desires and preferences tend to be shaped inevitably and that this justifies nudging because there is usually no autonomy to be interfered with, see the text following footnote 72 below.

This would not be such a great worry if there was some objective welfare to be pursued without the possibility of contention. Such a scenario, however, may be less realisable than first appears. Thus, a nudge might aim to further ‘better health’ on the assumption that this is uncontroversial but, as Gregory Mitchell has pointed out: ‘Many may agree in the abstract that better health is preferable to worse health, but when the choice is framed as enjoying life-shortening but intensely pleasurable vices during one’s college days versus abstaining during college to gain a couple of extra boring years at an advanced age, then better health may not look quite as good.’⁶¹

The danger is that the enthusiastic proponents of nudging⁶² will be slow to anticipate conceptions of welfare or preferences that they see as non-sensible or which stem from perspectives that they do not share. They may, moreover, fail to anticipate that nudges may prove contentious not merely because they seek to further contested visions of welfare but because the costs of furthering those visions, or the distributions of such costs, or the costs of opting out of nudges may be controversial matters in themselves. A further concern of nudge’s critics may be that the banner of libertarian paternalism may be used as a cover for the pursuit of social objectives (such as lowering hospital’s administration costs) rather than the welfares of the nudged individuals.⁶³

Proponents of nudge, of course, have a fall-back defence against the accusations of illiberalism that apply to Second and Third Degree nudges: they stress, as noted, that a nudge should be ‘easy and cheap to avoid’.⁶⁴ A difficulty here is that there is an inherent tension in the nudging argument that goes to the feasibility of opt-out. The nudge has power in so far as it impacts on the decision-maker of limited cognitive capacity, information and self-control. Those very limitations, however, mean that the target is unlikely to be well-placed to exercise the opt-out in cases of Second Degree nudges. With Third Degree nudges, the position is worse because the target will be ‘blocked’ from resorting to the opt-out. When a nudge is used to shape the target’s preferences (as when a vivid example of a harm is used to trigger aversion) the opt-out is taken out of play. As Mitchell argues, more generally: ‘If choice is as ‘sticky’ as Thaler and Sunstein claim, then people who lack the means to determine their true preferences in a given setting cannot make any real use of the opt-out provision and will simply stick with the default set by the libertarian paternalist. For these people libertarian paternalism is just paternalism.’⁶⁵ For such reasons, nudges have been said to be, on their face, as threatening to liberty, broadly understood, as is overt coercion.⁶⁶

The opt-out defence is also vulnerable to the criticism that it not only assumes a level of competence, rationality and volitional control that contradicts the underpinning assumptions of behavioural economics, but it understates the extent to which opt-outs discriminate against parties who are less able to exercise them. An opt-out that is a mere mouse click is easy enough to most persons but not to the person who does not possess a computer. An organ donation opt-out that is recorded

⁶¹ See G. Mitchell, ‘Libertarian Paternalism is an Oxymoron’ (2005) 99 *N. Western U. Law Review* 1245, 1268.

⁶² On nudge as ‘never knowingly undersold’ see R. Baldwin, *The New Scholarship: Celebrating the ‘I’ in Ideas* LSE Law, Society and Economy Working Papers 5/2012. Available at: http://www.lse.ac.uk/collections/law/wps/WPS2012-05_Baldwin.pdf

⁶³ Blumenthal-Barby and H. Burroughs, n 22 above, 4.

⁶⁴ Thaler and Sunstein, 2008, n 2 above, 6.

⁶⁵ Mitchell, n 61 above, 1254.

⁶⁶ See Hausman and Welch, n 14 above, 130.

on a driving licence would not be very accessible to a committed cyclist, and some smokers with mobility issues may not think that the 100 metre journey to the office smoking zone is an easy opt out from the design nudge of the departmental architect. In the medical field, it has been pointed out that written opt-outs are not that useful to linguistic minorities, and illiterate persons.⁶⁷

A further concern is that, even if exercising one opt-out may be relatively easy, cumulations of nudges (of the Second and Third Degree⁶⁸) produce an aggregate burden that will weigh down persons who would seek to negotiate their own way through everyday decision shapings so as to pursue their own aims and objectives. In a world of nudge, the worry is, the individual fights an unending battle to construct his or her own preferences and to play the opt-out game in order to further those preferences. Their autonomy, they may feel, dies the death of a thousand nudges.

To construct a simple example, let us suppose that supermarkets agree with the Government that on-line shoppers will be presented with default purchases that encourage healthy lifestyles and that the retailers in question nudge with enthusiasm. The shopper may avoid the default acquisition of fresh apples on supermarket X's website by a mouse click but then has to click to avoid purchasing broccoli, cod liver oil tablets, aspirins, glucosamate pills and thirty other products. This arrangement raises a number of issues. In the first instance, there is a question of how a nudge is conceived for the purposes of opt-out. The 'apples nudge' is avoided by a single click but the 'healthy lifestyle' nudge requires over thirty clicks to neutralise – which is likely to prove wearisome if the process afflicts all on-line shopping episodes. On one view, each nudge is avoided easily and complies with Thaler and Sunstein's conditions for nudging. On another view, the cumulative burden of opting out of this package of defaults is unacceptable.

A further worry about Second and Third degree nudging is prompted by the example under consideration. Where a Government uses agents, such as retailers, to implement such nudges this creates a specific set of problems.⁶⁹ Not only are there issues of consistency of interpretation, transparency and accountability, but there are challenges to the opt-out. Where different supermarkets implement the nudge in different ways, the consumer has to pay special attention to each opt-out because some nudgers' approaches and some 'nudged' products will be more questionable than others. (Will the dosages of glucosamine sulphate sold by supermarket Y's default actually bring benefits that justify the costs? Is the cardiac case for a daily aspirin made out?) For consumers, the effort required to address these issues only adds to the burdens of opting-out. They may resent having to negotiate their way through a brand-specific lifestyle training session every time they order some groceries.

Nor can it be assumed that the consequences of a failure to opt-out will be uniform across populations. Thus, for example, presumed consents for HIV testing may prove especially threatening to individuals who are liable to a range of particularly negative consequences if they fail the test – such as domestic violence or

⁶⁷ See M-A Jacob, 'Another look at the presumed-versus-informed consent dichotomy in post-mortem organ procurement.' (2006) 20(6) *Bioethics* 293-300.

⁶⁸ This point applies to a lesser extent to First Degree nudges since even warnings and simple items of information may cumulate to a mass that is seen as oppressive by the targets.

⁶⁹ Agency issues, of course, add to dangers that nudges will cause unintended consequences and this further undermines the argument that nudges are more easily targeted than commands.

job loss.⁷⁰ Where there are such divergencies of consequence, and these are not taken on board, the danger is that calculations about the balance between the value of the nudge and the costs of opt-out will produce nudges that discriminate against vulnerable parties.

A broader issue of fairness also arises with Second and Third Degree nudges – they impose costs on sensible persons in order to enhance the welfare of others who behave irresponsibly. To take an example, let us suppose that a government is concerned to control excessive alcohol consumption. It passes a law to require all retail outlets to place alcohol in locked cabinets, out of sight of shoppers. The (Second Degree) nudge derives from the conscious effort that is required in order for the shopper to obtain the alcohol (by specific request for service) and from the neutralising of any sales-enhancing layouts by the retailer. Leaving aside whether this nudge would limit consumption effectively, it can be noted that the nudge involves an inconvenience for, amongst others, the 93 per cent of the mature male, and 96 per cent of the mature female, population who can limit their consumption to below ‘harmful’ levels when left to their own devices and exposed to the nudges of retailers, brewers and distillers.⁷¹ There is, accordingly an issue of distributional justice here and some parties will object to this redistribution from the rational/responsible to the irrational/irresponsible.⁷²

A second defence of nudging is of special relevance to Second and Third Degree nudges and argues that there is always a nudge from some direction so that peoples’ preferences and choices are inevitably structured and influenced (either intentionally or unintentionally) and that, therefore, it is best that the nudge is applied in a welfare enhancing direction.⁷³ Thus, it might be contended that, uncontrolled, the astute supermarket would be likely to place a profit maximising product next to the till, or the less organised supermarket might rotate products randomly on the till display. Far better, it would be said, if the supermarket was required to place the fruit next to the till in order to nudge consumers towards healthy options.

A first objection to this defence is that it is wrong to assume that all preferences and choices are framed or presented in unacceptable ways. In many decision-making contexts, the decision-maker is capable of identifying his or her own preferences and is faced with information that he or she is quite capable of processing and acting on so as to serve their interests and preferences.⁷⁴ The real issue, on this view, is not whether the decision is shaped but whether the pre-nudge decision-making context is seen as manageable and acceptable. Even when nudged by, say, a retailer, it is routine for the consumer to be content with the rationality and level of autonomy that he or she exercises. The decision environment may not be ‘true’ but it is seen as acceptably manageable. (The consumer exercises his or her autonomy here in accepting that they, rather than someone else, will deal with the decision-making context in all its imperfections.) Thus, many vegetable purchasers know what they are looking for and

⁷⁰ C. Hanssens, ‘Legal and Ethical Implications of Opt-out HIV Testing.’ (2007) 45 *Clinical Infectious Diseases* 232-239.

⁷¹ In the 45-64 year old grouping, the evidence is that 7% of men and 4% of women drink to a ‘harmful’ level (over 50 units per week for men and 35 units for women) – see Institute of Alcohol Studies, *Alcohol Consumption Factsheet* (IAS, 2013) 11.

⁷² See Mitchell, n 61 above, 1270.

⁷³ First Degree nudges arguably do not demand that this defence is made since they enhance conscious decision-making.

⁷⁴ See the concession of this point at Thaler and Sunstein 2008, n 2 above, 10.

expect the greengrocer to place the best produce at the front of the stall and to serve from the back.

The advocates of Second and Third degree nudging, it can thus be argued, interfere with the autonomy of individuals too readily and nudge theory can be said not yet to have produced either a way of separating out unacceptable from acceptable choice structurings or of identifying those circumstances in which decision-making autonomy is so prejudiced by shaping influences that this justifies intervention through another nudge.⁷⁵

A closely related line of nudge-defence should be considered at this point. This contends that it is wrong to object to Second and Third degree nudging as an attack on individual autonomy because such an objection presupposes an idealised version of ‘true’ autonomy that is based on an assumption of fully-informed, rational and unshaped decision-making. The reality, the nudge proponents argue, is that the ‘autonomous choices’ that individuals generally make are usually shaped by a host of features of the consumption environment, they are often the products of cognitive and volitional limitations, and, as a result, there is rarely an exercise of unimpaired and ‘true’ autonomy. The counter to this argument, however, is not merely the point just made (that some choice framings are seen as acceptable) it is also that the pro-nudgers are too quick to portray some preferences as irrationalities. Thus, some of the ‘biases and blunders’ that Thaler and Sunstein cite as causes of poor decisions⁷⁶ can be said to be preferences that deserve to be respected rather than cognitive or volitional failings that need to be reacted to with a nudge. Loss aversion, for instance, is said by Thaler and Sunstein to be a ‘matter that people can get pretty emotional about’⁷⁷ so that they value not losing things far more than gaining the same things. This produces inertia and leads to decisions that do not serve peoples’ own interests.⁷⁸ It is a ‘fallibility’ that the two authors suggest can be corrected with a nudge. The nudge-sceptics, however, would argue that this is not a case of impaired autonomy but of a choice that the nudgers disagree with. The sceptics might add that, if a person has a preference for the status quo, it is their life, their inertia, their choice, not that of the nudger.

To return to reservations about the ‘nudges are unavoidable’ argument, a second objection is that, even if the premise is assumed to be correct, the conclusion does not follow. Thus, it can be argued that the supermarkets’ propensities to nudge consumers in unhealthy directions (leading, for example, to the consumption of excessive amounts of sugar) do not necessarily demand that the state responds to such nudging with another nudge (or, indeed, with a nudge of a Second or Third degree as opposed to a First Degree). In many contexts, it could be argued, it would be more effective, more accountable and more transparent to react with, say, a command approach and revise the laws on food ingredients (e.g. to limit the maximum concentrations of sugar that are deemed safe to sell).

There is, indeed, an argument that to respond to nudges with state-applied counter nudges of the Second and Third Degree is exactly the wrong policy in those cases where it causes citizens and consumers to be exposed to layers of competing nudges.

⁷⁵ See Mitchell, n 61 above, 1249-50. Thaler and Sunstein do, however, offer a discussion of the circumstances in which people are most, and least, likely to make good decisions – see Thaler and Sunstein 2008, n 2 above, Chapter 4.

⁷⁶ Thaler and Sunstein 2008, n 2 above, Chapter 1.

⁷⁷ It is tempting, at this point, to hear the late Michael Winner saying: ‘Now calm down dear.’

⁷⁸ Thaler and Sunstein 2008, n 2 above, 37.

This accumulating of nudges, it could be argued, multiplies the undermining of autonomy that even single (Second or Third Degree) nudges involve.

Here we see that a supposed argument for nudging may have force in relation to one degree of nudge but not to others. Thus, when Thaler and Sunstein argue that ‘rare, difficult choices are good candidates for nudge’⁷⁹ it may be conceded that some positive potential is offered by a First Degree nudge that is designed to enhance conscious decision-making by rendering complex data more manageable. It will be far less readily acknowledged that there is case for a Second or Third Degree nudge since, as just argued, this is likely to conduce to a cumulative undermining of autonomy.

A further concern about Second and Third Degree nudges relates to their longer-term and extensive use. It is that such responses do not seek to improve rational, informed decision-making but serve to reduce the citizen’s voice in his or her destiny. Used over time and numbers of issues, the effects of such accumulations of nudging may be to produce control regimes that are defeatist about the capacities of individuals to become more responsible and deliberative. The worry is that proliferations of nudges produce citizens with ‘choice structuring fatigue’ and they positively discourage the kinds of learning, responsible and thinking behaviour that most governments espouse. Any short term positive gains from nudge may be outweighed in time as people both ‘zone out’ from nudges and succumb to an ‘infantilisation effect’ by assuming less responsibility for their own welfare.⁸⁰

It can be argued, furthermore, that many citizens may be unhappy about *state* Second and Third Degree nudges in a way that they are not about commercial or other decision-shaping pushes.⁸¹ There may be a number of reasons for this. They may see commercial nudges, such as advertising campaigns and store designs, as consistent and predictable facts of life that they are comfortable in dealing with. They accept that salespersons will puff their products and engage in like behaviour but they may even enjoy negotiating their way through the market, provided that no lies are told. In contrast, they may see state nudges of the Second and Third Degree to be far more covert, unknowable and threatening than the expected strategies of marketing and sales. The public, moreover, may fear the scale of organized governmental nudging much more than they would a random collection of market-driven nudges.

Expectations may also differentiate between nudges from private and public sources. Citizens and consumers may anticipate and accept that the private sector will behave in a profit-enhancing manner and use methods to further such ends that are not wholly transparent. The public, though, may expect better things from the governments that they have elected and entrusted with vast sets of powers. There is likely to be a demand that, in using these powers to control and shape peoples’ lives

⁷⁹ Thaler and Sunstein 2008, n 2 above, 82.

⁸⁰ See Bovens n 14 above, 11. For similar concerns that ‘design’ or ‘code’ approaches to regulation fail to respect citizens as morally responsible agents see Brownsword n 52 above 17-20 and D. Smith, ‘Changing Situations and Changing People’ in A. Von Hirsch, D. Garland, A. Wakefield (eds.), *Ethical and Social Perspectives on Situational Crime Prevention* (Oxford: Hart Publishing, 2000) 147-174 (who notes the ‘demoralising effect’ of regimes that remove personal choice).

⁸¹ Thaler and Sunstein note this worry (n 2 above, 11) but downplay it on the grounds that nudges are more benign than commands and because nudges are inevitable. In portraying this concern as a ‘non-starter’ Thaler and Sunstein do not take the point that nudges can be far less transparent than commands. The ‘inevitability of the nudge’ contention also sidesteps the point that nudges from some sources may be more acceptable than from others.

and decisions, ministers, officials and regulators use methods that are properly open and accountable rather than covert and undermining of citizen autonomy. Nudges of the second and third degree are liable to be seen as exploitative of citizens' weaknesses and as 'a form of disrespectful social control'.⁸² The public, on some issues, will think: 'If the government sees the need to control this issue, they should tell us this and have the courage to seek parliamentary approval for new legislation rather than operate with some sly nudge.'

A third defence of nudging would contend that the allegedly negative features of nudge have to be weighed against the behavioural and welfare gains that can be achieved. This is a defence that is canvassed by some commentators but it is not one that Thaler and Sunstein rely on heavily, since they take it that the heat is taken out of this issue by their stipulation that opt-outs should be easy.⁸³

To balance the costs and benefits of nudging is, however, just one of a number of ways of dealing with the case for nudging and two further approaches can be noted. A second considers the proportionality of the nudge and a third looks to the public acceptability of the nudge.

The essence of the cost-benefit approach is a comparison of the gains in health or other aspects of welfare to be anticipated from a nudge and the losses to autonomy and quality of life that a nudge may produce in the short or longer term.⁸⁴ A 'proportionality' approach differs from a cost-benefit view in emphasising non-quantitative and ethical factors. It adverts to such matters as the scale of the problem to be addressed, the evidence that the nudge will be effective, and the representative and moral considerations that are relevant. These latter will include such matters as intrusiveness, restriction of freedom, quality of choices made, and transparency.⁸⁵

A public acceptability approach asks whether the use of a nudge, in a given context, is sanctioned by popular approval or 'public permission'.⁸⁶ For its part, the House of Lords Science and Technology Select Committee was not convinced by this latter approach and cautioned that: the public tended to resist change and its approval might only be evident sometime after a nudge had been used. Popular approval tended to vacillate over time, and the individuals making up the public might possess distorted or misguided views on these matters for a number of reasons familiar to behavioural economists.⁸⁷

In the following section the approach that is adopted in assessing nudge comes closest to the proportionality version. It draws out the messages from the above discussion in seeking to identify the circumstances under which governments can be expected to use nudge in a manner that is positive.

⁸² Hausman and Welch, n 14 above, 134.

⁸³ *Ibid.* See, however, Thaler and Sunstein 2008, n 2 above, 249 where the authors admit that they have no 'inflexible rule' on when opt-out costs are excessive and argue that: 'The real question is when we should be willing to impose some nontrivial costs in the interests of improving people's welfare.'

⁸⁴ The message of the behavioural approach that preferences are shaped, of course, renders calculations of costs and benefits fraught because this calls into question the primary basis for defining and assessing welfare – namely the satisfaction of individual preferences. (I am grateful to Andrew Lang for making this point.)

⁸⁵ See House of Lords Select Committee, 2011 para 2.19.

⁸⁶ See House of Lords Select Committee, 2011 para 2.16.

⁸⁷ *ibid.*

THE CONDITIONS FOR EFFECTIVE AND ACCEPTABLE USE

It is clear that First, Second and Third Degree nudges raise very different issues and that this has to be borne in mind in looking at any potential nudge. Not only that but different nudging tools have divergent potential in different contexts. This section will, accordingly seek to deal with the different degrees of nudge and the different tools for nudging whenever these present variations of challenge.

Can the conditions for effective use of nudge be stated? A first message from the above discussion is that nudge cannot be expected to produce results effectively, as a sole strategy for intervention, where the causes of the identified mischief go beyond individual choice and relate to such matters as economic and cultural environments. Secondly, it is a condition for effective use of nudge that the targets of the given nudge are liable to be affected by the particular nudge tool employed. Table 2 indicated how dramatically reactions to nudges can be expected to vary across not only targets with different characteristics, but also tool types, degrees of nudge, and modes of implementing nudge tools.

Third, unanticipated side effects and unintended consequences have to be anticipated as much with nudges as command rules, taxation mechanisms or other intervention strategies. Similarly, it is necessary to anticipate the effectiveness implications of any likely counter-nudges that may be forthcoming.

Fourth, mechanisms have to be adopted to deal with the danger that nudging objectives lose contact with the relevant expectations of the public or the government. This may demand that deliberative procedures are operated in parallel with nudges and that feedback loops are established so as to inform and update the nudgers.

Fifth, the cumulative and long-term effects of nudging have to be taken on board in order to avoid the tendency of short term gains to be achieved and the expense of longer term, and greater, costs. Such effects include potential diminutions in trust, nudge fatigue, and infantilisation.

Finally, a condition for the effective use of nudges across government is that governments institute, on a broad basis, more studies that are designed to evaluate behaviour change interventions.

Turning to the conditions for using a proposed nudge in an acceptable manner, the central challenge is to satisfy expectations on representational and ethical fronts. Regarding such matters, First Degree nudges can be separated out as far less problematic than Second or Third Degree nudges. With a First Degree nudge, the autonomy of the citizen is respected and conscious decision-making is enhanced. Few representational or ethical problems can be expected to arise provided that the simple information (or reminder) given is accurate. Some attention, nevertheless, may have to be paid to the need to justify the decision to provide this piece of information or this reminder rather than another.

Second Degree and Third Degree nudges are a different matter because they detract from the autonomy of individuals to different extents. Acceptable use of nudge demands that the nudgers have satisfactory responses to the following concerns. The lack of public debate and accountability involved in many nudges. The imposition of the nudger's conception of welfare rather than the target's. The limitations of opt-outs, their variations in availability and practicality and their sometimes discriminatory effects. A further concern that may call for a response is the fairness of imposing on the autonomy and convenience of more sensible citizens in order to assist those who are less responsible. With Third Degree nudges, the emotional/psychological blocking of the opt-out is a special worry as is the removal of opt-out through the re-shaping of preferences.

Justification may also be needed for choosing the nudge as a response to a problem (even a decision-framing problem) rather than a command, economic incentive or other intervention. The notion that nudges are inevitable, and that this justifies governmental nudging, seems flawed on a number of counts.

How then, might one go about judging whether a proposed nudge would have value in a given context? A starting point would be clarity on the character of the mooted nudge. Is it a First, Second, or Third Degree example? Such characterisation immediately throws challenges into relief. Next, it is essential to demonstrate with an evidentially secure case that the outcomes of the proposed nudge are likely to be secured effectively and that nudging will realise these in a superior fashion to more traditional regulatory interventions.

On representational and ethical matters, one potential response to worries is to show that processes ancillary to the nudge will ameliorate problems. Thus, references can be made to transparency exercises such as disclosures of nudges or consultation and feedback procedures. It can be expected, however, that, as noted above, general strategies of disclosure will be seen as doing little to change the semi-covert nature of many Second Degree nudges, and that disclosures relating to Third Degree nudges are as likely to provoke protest as to reassure potentially targeted citizens.

A second response is to demonstrate that there are no representational and ethical losses resulting from nudging since, in the given context, there is evidence that individual autonomy is already so badly compromised that a nudge produces a welfare enhancing result without further detracting from decision-making autonomy. A difficulty here is to show that the alleged imperfections of autonomy that are cited are not, in fact, merely manifestations of preference that the nudger lacks sympathy with. A further problem with the 'compromised autonomy' argument lies in contending that Second and Third Degree nudges will enhance decision-making rather than simply place a further hurdle in the way of conscious decision-making.

Overall, the case has to be made that the gains from nudging will outweigh the representational and ethical costs and worries noted above. The difficulty is that Second and Third Degree nudges will routinely involve representational and ethical costs that are both difficult to quantify and liable to be contested. A further complication is that judgements will often have to be made about the acceptable balance between short term gains and longer term losses.

In short, it is easier to identify the challenges and concerns involved in using nudges than to quantify the prospects of making a good case for nudging in any single context or more generally as a tool of governmental influence.

ANOTHER TOOL IN THE BOX?

At first glance, the safest conclusion to draw about nudge would seem to be that this may not be the complete answer to governmental control but it is, at least, a very useful device to have in the regulatory / non-regulatory toolbox of influences over conduct. Further consideration, however, suggests that this conclusion builds too readily on a particular conception of the toolbox of state control devices: one in which different regulatory and non-regulatory tools take their places in an harmonious group (a sort of seven dwarves choir of governance).

The reality is often that different such tools may compete rather than co-ordinate harmoniously. In the first instance, the tools may vie for political prominence. Thus the House of Lords Science and Technology Select Committee expressed the concern in 2011 that the UK government's promotion of behaviour change strategies was

blocking out consideration of other intervention strategies.⁸⁸ This typifies the worry that when politicians and ministers fall in love with a certain newly fashionable policy style or mode of intervention, they tend to be blind to the virtues of the more traditional approaches. This is a danger that can be responded to (at least partially) with a First Degree nudge – as exemplified by the Science and Technology Committee issuing a reminder to the Government that it should not forget that a range of policy tools is available.

There is, moreover, a second more structural worry about seeing control tools as an essentially harmonious package of options. Different regulatory and non-regulatory tools of intervention operate with different logics in so far as they establish different relationships with targeted persons and use varying ways of making demands of such targets.⁸⁹ The covert aspect of Second and Third Degree nudges involves little dialogue between the nudger and the target. Indeed, as noted above, an openness of dialogue can be expected to remove much of the force from certain types of nudge. This contrasts with the position in more traditional regulation when steps are taken routinely to clarify requirements, explain what compliance demands and so on. The problem in combining intervention tools such as commands and nudges is that they may undermine each other – it is difficult to establish a control relationship that is based on openness, and, at the same time, to use nudges by stealth. The effect of introducing nudging into an area may, thus, be to reduce the effectiveness and the transparency of other modes of intervention and to detract from the broader regime of control. This point was made above in noting that a problem of using nudges in the medical field is that this can destroy the trust relationships that underpin the other interactions between patients and physicians.⁹⁰

A third problem in seeing the nudge as just another tool in the box is that choices of different intervention tools are governed by different philosophies of control and of state/individual relationships. It has been seen above that nudging operates on the basis of libertarian paternalism so that flawed individuals' decisions stand to be corrected by the state so as to increase citizens' welfare. With Second and Third Degree nudges, the basis of the nudge is an exploitation of the limited cognitive, volitional and emotional capacities of individuals. The idea in such paternalism, is not to take steps in order to increase the quality of individuals' rational, conscious decision-making, it is to accept peoples' flaws and use these to advantage.

The difficulty regarding the toolbox is that some tools of intervention operate on quite different philosophical and ethical assumptions from those that underpin nudge.⁹¹ Classical command and control rules of regulation, and economic incentives, tend to build on the rationality of persons and firms – they issue prescriptions that are overt because they take it that people will use their rationality to react positively to the threats of sanctions or taxes. Enforcement practices tend to work on a rational deterrence basis, and models such as the 'responsive regulation' approach, treat regulated persons and firms as cognitively competent responders to overtly issued messages.

The implication of these differences is that nudge does not sit wholly happily alongside many other intervention tools. Choices have to be made about the position

⁸⁸ House of Lords Select Committee, 2011 para 8.15.

⁸⁹ See R. Baldwin and J. Black, 'Really Responsive Regulation' (2008) 71 *Modern Law Review* 59-94.

⁹⁰ Blumenthal-Barby and H. Burroughs, n 22 above, 5.

⁹¹ See the parallels with Brownsword, above n 52 on the differences between 'East coast' and 'West coast' intervention styles.

that the state wants to adopt regarding targeted individuals – is their autonomy to be respected and their rationality to be openly worked with, or is their autonomy to be reduced and their decision-making to be nudged by stealth? A decision by a government to adopt one of these philosophies rather than the other involves a disposition for or against using nudges as opposed to more traditional methods. For this reason, a vision of the nudge as just another potentially useful intervention tool may be one that is too easily arrived at.

This is not to argue that nudges have no role within government, it is to caution that, when choosing when to nudge rather than intervene in another way, it is essential to be clear and open about the philosophical basis for such a choice as well as to be aware that different modes of intervention may operate with clashes of logic that threaten not only effectiveness but also the serving of representative and ethical ends.

CONCLUSIONS

The nudge has burst on the scene in the last decade and it has proved enormously alluring to certain governments, including those of the UK and USA. There are, however, dangers in embracing the nudge uncritically. A first problem has been that the nudge has often been discussed, at least in policy circles, with a lack of conceptual precision. This article has sought to address this deficiency by distinguishing between three distinct degrees of nudge and by exploring the different concerns that relate to nudges of these three orders. This, at least, ensures that the case for nudging in a given context will not be assessed on the assumption that all behaviour change strategies are as benign as the First Degree nudges described here.

It has been seen that it is difficult to state, in precise terms, the conditions under which nudging can be used effectively and acceptably. This is because there are numbers of different nudging tools and three different degrees of nudging severity to be taken into account alongside varieties of target. It has been possible, however, to identify the challenges and concerns that are posed by nudges of different degrees and to assess the extent to which these can be responded to in positive terms.

It should be emphasised, finally, that the above parade of reservations about nudges does not prompt the conclusion that the nudge should be abandoned. The argument here is that the nudge strategy must be used with precision and an awareness of both its limitations and its fit within the range of state interventions. The three degrees of nudge are here to stay and there will be more instances of behaviour change – one can only ask of the nudge: ‘When will I see you again?’