

ISSN 1936-5349 (print)
ISSN 1936-5357 (online)

HARVARD

JOHN M. OLIN CENTER FOR LAW, ECONOMICS, AND BUSINESS

NUDGING AND CHOICE ARCHITECTURE: ETHICAL CONSIDERATIONS

Cass R. Sunstein

Forthcoming in *Yale Journal on Regulation*

Discussion Paper No. 809

01/2015

Harvard Law School
Cambridge, MA 02138

This paper can be downloaded without charge from:

The Harvard John M. Olin Discussion Paper Series:
http://www.law.harvard.edu/programs/olin_center/

The Social Science Research Network Electronic Paper Collection:
<http://ssrn.com/abstract=2551264>

Nudging and Choice Architecture: Ethical Considerations

Cass R. Sunstein *

Abstract

Is nudging unethical? Is choice architecture a problem for a free society? This essay defends seven propositions:

(1) It is pointless to object to choice architecture or nudging as such. Choice architecture cannot be avoided. Nature itself nudges; so does the weather; so do customs and traditions; so do spontaneous orders and invisible hands. The private sector inevitably nudges, as does the government. It is reasonable to worry about nudges by government and to object to particular nudges, but not to nudging in general.

(2) In this context, ethical abstractions (for example, about autonomy, dignity, manipulation, and democratic self-government) can create serious confusion. To make progress, those abstractions must be brought into contact with concrete practices. Nudging and choice architecture take highly diverse forms, and the force of an ethical objection depends on the specific form.

(3) If welfare is our guide, much nudging is actually required on ethical grounds, even if it comes from government.

(4) If autonomy is our guide, much nudging is also required on ethical grounds, in part because some nudges actually promote autonomy, in part because some nudges enable people to devote their limited time and attention to their most important concerns.

(5) Choice architecture should not, and need not, compromise either dignity or self-government, but it is important to see that imaginable forms could do both. It follows that when they come from government, choice architecture and nudges should not be immune from a burden of justification, which they might not be able to overcome.

(6) Some nudges are objectionable because the choice architect has illicit ends. When the ends are legitimate, and when nudges are fully transparent and subject to public scrutiny, a convincing ethical objection is less likely to be available.

(7) There is ample room for ethical objections in the case of well-motivated but manipulative interventions, certainly if people have not consented to them; such nudges can undermine autonomy and dignity. It follows that both the concept and

* Robert Walmsley University Professor, Harvard University. This essay was the basis for remarks on the ethics of choice architecture and nudging, to be delivered at a conference on that topic at Humboldt University in Berlin in January 2015. I am most grateful to participants in the conference for many valuable thoughts and suggestions. I am also most grateful to Matthew Lipka, Martha Nussbaum, Lucia Reisch, and Adrian Vermeule for superb comments on a previous draft. This essay is a substantial revision and expansion of an earlier one, *The Ethics of Nudging*, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2526341.

the practice of manipulation deserve careful attention. The concept of manipulation has a core and a periphery; some interventions fit within the core, others within the periphery, and others outside of both.

Table of Contents

- I. The Central Argument
- II. Concepts and Definitions
 - A. In General
 - B. The Inevitability of Choice Architecture
 - C. Spontaneous Orders and Visible Hands
- III. The Diversity of Nudges and the Trap of Abstraction
 - A. Motivating Nudges
 - B. Three Distinctions
 - C. Illicit Reasons and Transparency
- IV. “As Judged By Themselves”
 - A. The Basic Standard
 - B. Questions and Doubts
 - 1. Objectively good lives?
 - 2. Ex ante or ex post
 - 3. Preferences about preferences
 - 4. Informed judgments and empirical puzzles
 - 5. Self-control problems
 - C. Bounded Rationality: Education and its Discontents
- V. Seven Objections
 - A. Paternalism
 - B. Autonomy
 - C. Coercion
 - D. Dignity
 - E. Manipulation
 - 1. Manipulation and Conscious Choice
 - 2. Manipulation, Self-Interest, and the Man Behind the Curtain
 - A. Learning
 - B. Biased Officials
- VI. What Do People Think? Some Evidence
 - A. The Effects of Transparency About Nudging
 - B. Politics Matters
 - C. Nudging System 1
- VII. Conclusion

I. The Central Argument

1. The goal of this Essay is to explore the principal ethical objections to nudges and choice architecture.¹ I aim to provide a taxonomy of those objections and to assess the force of each.
2. The last decade has seen a remarkably rapid growth of interest in choice-preserving, low-cost regulatory tools.² Especially in light of that interest, it is important to obtain an understanding of the nature and weight of the ethical concerns. Many regulatory tools, and perhaps increasingly many, involve nudges and choice architecture, and some of the ethical objections can have either political or legal resonance. For example, nudges that involve information disclosure, or compulsory warnings, might raise first amendment issues.
3. My central argument is that at least if they are taken in general or in the abstract, the ethical objections lack much force, and for two different reasons. *First*, both nudges and choice architecture are inevitable, and it is therefore pointless to wish them away. *Second*, many nudges, and many forms of choice architecture, are defensible and even required on ethical grounds, whether we care about welfare, autonomy, dignity, self-government, fair distribution, or some other value.
4. It is true that all government action, including nudges, should face a burden of justification (and sometimes a heavy burden). If the government requires disclosure of information, or establishes particular default rules, it must explain and defend itself. The fact that people retain freedom of choice, and are ultimately permitted to go their own way, does not give public officials a kind of license to do whatever they want.³ But in many cases, the requisite explanation is available.
5. Suppose, for example, that we are welfarists and hence believe that the goal of social ordering (including those forms for which government is responsible) is to promote social welfare. If so, we will favor welfare-promoting nudges. Or suppose that we believe in individual autonomy and dignity. If so, we will favor nudges and choice architecture that

¹ Some of strongest objections can be found in Riccardo Rebonato, *Taking Liberties* (201).

² Catalogues can be found in OECD, *Regulatory Policy and Behavioral Economics* (2014); Cass R. Sunstein, *A Council of Psychological Advisers?*, *Annual Review of Psychology* (forthcoming 2015).

³ Note as well that a disclosure requirement is a mandate, and no mere nudge, for the people on whom the requirement is imposed. It might be a nudge for consumers but a requirement for producers. I will say more about this point below.

promote those values. (It is possible, of course, that distrust of government, and faith in markets, will incline us to minimize nudging on welfarist or autonomy grounds.⁴)

6. If we value democratic self-government, we will be inclined to support nudges and choice architecture that can claim a democratic pedigree and that promote democratic goals. Any democracy has a form of choice architecture that helps define and constitute its own aspirations to self-government. A Constitution can be seen as a kind of choice architecture for choice architects. A self-governing society might well nudge its citizens to participate in the political process and to vote. (Certain political parties engage in such nudging.) Collective precommitment strategies, designed for example to promote public health, might produce nudges (or even mandates), and they might be justified on democratic grounds.⁵
7. Of course no one should approve of nudges or choice architecture in the abstract or as such. Some nudges, and some forms of choice architecture, do indeed run into convincing ethical objections. Suppose, for example, that a nation establishes a default rule stating that unless voters explicitly indicate otherwise, they will be presumed to support the incumbent leader in the election. Or suppose that a nation establishes a default rule to the effect that unless citizens indicate otherwise, their estates will revert to the nation's most powerful political party upon their death. There is ample reason to question a default rule of this kind even if citizens are authorized to opt out.
8. A central question is whether nudges and choice architecture promote welfare or autonomy and dignity. Another question is whether they are consistent with democratic norms. Some nudges have illicit ends, and they are objectionable for that reason.
9. There is also a pervasive question about manipulation. As we shall see, transparency and accountability are indispensable safeguards, and both nudges and choice architecture should be transparent.⁶ Even if so, there is a risk of manipulation, and that risk should be avoided. Many of the

⁴ Edward L. Glaeser, *Paternalism and Psychology*, 73 U. Chi. L. Rev. 133 (2006)

⁵ For discussion, see Jon Elster, *Ulysses and the Sirens* (1983); Geoffrey Brennan and L. LOMASKY (1983), *Institutional Aspects of Merit Goods Analysis*, *Finanzarchiv* N.F. 41, pp. 183 – 206. In the particular context here, see GEBHARD KIRCHGÄSSNER, *Soft Paternalism, Merit Goods, and Normative Individualism* (2014), available at https://www.wiwi.tu-clausthal.de/fileadmin/Volkswirtschaftslehre/VfS-Wipol_Ausschuss/Kirchgaessner_Soft_Paternalism.pdf

⁶ Of course the term needs definition. At a minimum, I mean to suggest that no nudging should be hidden or free from public scrutiny. See below.

most interesting and complex ethical questions involve the disputed concept of manipulation, and I will devote some attention to that concept here.

10. More specifically, I will offer seven principal conclusions.

(a) It is pointless to object to choice architecture or nudging as such. The private sector inevitably nudges, as does the government. We can object to particular nudges, and particular goals of choice architects, but not to nudging in general. For human beings (or for that matter dogs and cats and mice), choice architecture cannot be avoided. It is tempting to defend nudging on the part of government by saying that the private sector already nudges (sometimes selfishly) – but this defense is not necessary, because government is nudging even if it does not want to do so.

(b) In this context, ethical abstractions (about, for example, autonomy, dignity, and manipulation⁷) can create serious confusion. We need to bring those abstractions into contact with concrete practices. Nudging takes many diverse forms, and the force of an ethical objection depends on the specific form.⁸

(c) If welfare is our guide, much nudging is actually required on ethical grounds, even if it comes from the government. A failure to nudge, and not merely the right kind of nudge, might be ethically problematic, at least if we do not insist on controversial (and possibly incoherent) distinctions between acts and omissions.⁹

(d) If autonomy is our guide, much nudging is also required on ethical grounds. Some nudges actually promote autonomy, by ensuring that choices are informed and that choices will actually be made. Some nudges promote autonomy by freeing people to focus on their real concerns; there is a close relationship between autonomy and time

⁷ A valuable discussion is T. M. Wilkinson, *Nudging and Manipulation*, 61 *Political Studies* 341 (2013). See also Sarah Conly, *Against Autonomy* 30 (2012): “Libertarian Paternalism is manipulative. That is, it does not suggest that we engage in free and open discussion in order to rationally persuade you to change your ways. . . . The point of the nudge is to push you in ways that bypass your reasoning. That is, they use your cognitive biases, like your tendency to go with the default option, to bring about good effects. There is a sense in which they fail to respect people’s decision-making ability.” The concern must be engaged in the context of (some) nudging.

⁸ *Id.*

⁹ On acts, omissions, and government, see Adrian Vermeule and Cass R. Sunstein, *Is Capital Punishment Morally Required? Acts, Omissions, and Life-Life Tradeoffs*, 58 *Stan. L. Rev.* 703 (2005).

management. A failure to nudge might seriously compromise autonomy.

(e) Choice architecture should not, and need not, compromise either dignity or self-government, though imaginable forms could do both. The value of dignity (explicitly recognized in the German Constitution¹⁰ and playing a significant role in American constitutional law as well) imposes a barrier to some forms of choice architecture and some nudges. Self-government calls for certain nudges, and legitimates others, and forbids still others.

(f) Many nudges are objectionable because the choice architect has illicit ends. If the ends are legitimate, and if nudges are fully transparent and subject to public scrutiny, a convincing ethical objection is less likely.

(g) There is nonetheless room for such an objection in the case of manipulative interventions, certainly if people have not consented to them. The concept of manipulation deserves careful attention, especially because manipulation takes many forms, and can compromise both autonomy and dignity. Some forms of manipulation are built into the fabric of everyday life, including relationships between friends and even spouses. Advertisements and storefronts manipulate. Nonetheless, manipulation can run into serious objections, perhaps especially when it comes from governments.

II. Concepts and Definitions

A. In General

11. Nudges are interventions that steer people in particular directions but that also allow them to go their own way.¹¹ A reminder is a nudge; so is a warning. A GPS nudges; a default rule nudges. To qualify as a nudge, an intervention must not impose significant material incentives.¹² A subsidy is not a nudge; a tax is not a nudge; a fine or a jail sentence is not a nudge. To count as such, a nudge must fully preserve freedom of choice. If an intervention imposes significant material costs on choosers, it might of course be justified, but it is not a nudge.¹³ Some nudges work because they inform people; other nudges work because

¹⁰ Article 1, paragraph 1 of the German Constitution states: “Human dignity shall be inviolable. To respect and protect it shall be the duty of all state authority.”

¹¹ See Richard Thaler and Cass R. Sunstein, *Nudge* (2008).

¹² On some of the complexities here, see Cass R. Sunstein, *Why Nudge?* (2014).

¹³ See *id.*

they make certain choice easier; still other nudges work because of the power of inertia and procrastination.

12. When people make decisions, they do so against a background consisting of choice architecture.¹⁴ A cafeteria has a design, and the design will affect what people choose. The same is true of websites. Department stores have architectures, and they can be designed so as to promote or discourage certain choices (such as leaving without making a purchase). Even if the layout of a department store is a result of chance, or does not reflect the slightest effort to steer people, it will likely have consequences on what people end up selecting.¹⁵ If people see certain items first, they are more likely to buy them.¹⁶
13. Both private and public institutions (include courts) create default rules. In fact they cannot dispense with them. A cell phone, a mortgage, a tablet, and a welfare program will inevitably come with defaults, which can be changed if the relevant people agree. The law of contract is permeated with default rules, which establish what happens if people do nothing. Default rules nudge.
14. Even if a default rule is chosen on the ground that it captures what most people will do, and is in that sense “market-mimicking,” it will likely have some effect on preferences and outcomes. A default establishes initial entitlements, and it can be important for that reason, influencing people’s preferences.¹⁷ That influence raises ethical issues of its own. For present purposes, the point is that default rules, of one or another kind, are sometimes unavoidable, or practically so.¹⁸
15. Attention is a scarce resource. When applications (for loans, for educational opportunities, for refinancing mortgages, for training, for financial benefits of any kind) are complex and difficult, people may not apply; a great deal of money might be lost as a result.¹⁹ This point has implications for regulatory design. It suggests that the private sector may help or hurt people by focusing their attention in certain ways. The same is true for the public sector, whether or not it seeks to do so. A

¹⁴ See *id.*

¹⁵ See Brian Wansink, *Slim By Design* (2014).

¹⁶ Eran Dayan and Maya Bar-Hillel, *Nudge to Nobesity II: Menu Positions Influence Food Orders*, 6 *Judgment and Decision Making* 333 (2011).

¹⁷ See Russell Korobkin, *The Endowment Effect and Legal Analysis*, 97 *Nw L Rev* 1227 (2003).

¹⁸ An alternative, of course, is to call for active choosing, but sometimes that is not feasible. See Cass R. Sunstein, *Choosing Not To Choose* (forthcoming 2015).

¹⁹ See Benjamin Keys et al., *Failure to Refinance* (2014), available at <http://www.nber.org/papers/w20401>

regulation might be written or applied in a way that makes certain features of a situation especially salient.

16. Spontaneous orders are celebrated by many people,²⁰ and very reasonably so. If an order is spontaneous, there is something reason to think that it reflects the judgments of many people about how it makes sense to proceed. On certain assumptions, spontaneous orders can promote people's welfare.²¹ but they are form of choice architecture no less than intentional designs, and they will include a measure of nudging, not least if they create and perpetuate social norms.²² Invisible hands can nudge every bit as much as the most visible ones. To be sure, spontaneous order and invisible hands may be less dangerous than intentional designs, and on certain assumptions²³ they are likely to be benign (or better); but they are nonetheless forms of choice architecture.
17. For the future, we could imagine new forms of choice architecture that are designed to improve antipoverty programs²⁴; environmental programs²⁵; energy programs²⁶; retirement and social security programs²⁷; anti-obesity programs²⁸; educational programs²⁹; health care programs; and programs to increase organ donation.³⁰ We could also imagine forms of choice architecture that are designed to combat race and sex discrimination,³¹ to help disabled people, and to promote

²⁰ See Friedrich Hayek, *Freedom, Reason, and Tradition* 229 (1958); Friedrich Hayek, *The Market and Other Orders* (2014).

²¹ See *id.*; Elinor Ostrom, *Governing the Commons* (1990).

²² See Edna Ullmann-Margalit, *The Emergence of Norms* (1976).

²³ Edna Ullmann-Margalit, *The Invisible Hand and the Cunning of Reason*, 64 *Social Research* 181 (1997).

²⁴ See Sendhil Mullainathan and Eldar Shafir, *Scarcity* (2013).

²⁵ Cass R. Sunstein and Lucia Reisch, *Automatically Green*, 38 *Harv Env L Rev* 128 (2014).

²⁶ *Id.*

²⁷ Ryan Bubb and Richard Pildes, *Why Behavioral Economics Trims its Sails*, *Harv L Rev* (2014).

²⁸ See Brian Wansink, *Slim By Design* (2014).

²⁹ See Adam Lavecchia et al., *Behavioral Economics of Education* (2014), available at <http://www.nber.org/papers/w20609>

³⁰ For an interesting empirical result, see Judd Kessler and Alvin Roth, *Don't Take 'No' For An Answer: An Experiment With Actual Organ Donor Registrations* (2014), available at <http://www.nber.org/papers/w20378> (finding that required active choosing has a smaller effect, in terms of getting people to sign up for organ donation, than prompted choice).

³¹ See Iris Bohnet et al., *When Performance Trumps Gender Bias: Joint Versus Separate Evaluation* (2013), available at

economic growth. A great deal of future work needs to be devoted to choice architecture in these and related domains.³²

18. There is no question that certain nudges, and certain kinds of choice architecture, can raise serious ethical problems.³³ Consider, for example, a government that used nudges to promote discrimination on the basis of race, sex, or religion. Any fascist government might well (and almost certainly does) nudge. Terrorists nudge. Even truthful information (for example, about crime rates) might fan the flames of violence and prejudice. (If people learn that crime is widespread, they might be more likely to engage in crime, because it is the social norm.³⁴) Groups or nations that are committed to violence often enlist nudges in their cause. Even if nudges do not have illicit ends, it is possible to wonder whether those who enlist them are treating people with respect.
19. The most prominent concerns about nudging and choice architecture point to four foundational commitments: (1) welfare, (2) autonomy, (3) dignity, and (4) self-government. Some nudges could run afoul of one or more of these commitments. It is easy to identify welfare-reducing nudges that lead people to waste time or money³⁵; an unhelpful default rule could fall in that category, as could an educational campaign designed to persuade people to purchase excessive insurance or to make foolish investments. Nudges could be, and often are, harmful to the environment.³⁶ Pollution is, in part, a product of unhelpful choice architecture.³⁷

B. The Inevitability of Choice Architecture

20. Consider in this light a tale from the novelist David Foster Wallace: “There are these two young fish swimming along and they happen to meet an older fish swimming the other way, who nods at them and says ‘Morning, boys. How’s the water?’ And the two young fish swim on for a bit, and then eventually one of them looks over at the other and goes

<http://www.montana.edu/nsfadvance/documents/PDFs/resources/WhenPerformanceTrumpsGenderBias.pdf>

³² See World Bank, World Development Report, *Mind and Society: How a Better Understanding of Human Behavior Can Improve Development Policy* (2015).

³³ See Mark White, *The Manipulation of Choice: Ethics and Libertarian Paternalism* (2013); Riccardo Rebonato, *Taking Liberties* (2011).

³⁴ See Thaler and Sunstein, *suora note*.

³⁵ See the weak effects of the opt-out design in Robert Letzler, *Knowing When To Quit: Default Choices, Demographics, and Fraud* (2014), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2512276

³⁶ For relevant discussion, see Sunstein and Reisch, *supra note*.

³⁷ See *id.*

‘What the hell is water?’”³⁸ This is a tale about choice architecture. Such architecture is inevitable, whether or not we see it. It is the equivalent of water. Weather is itself a form of choice architecture, because it influences what people decide.³⁹ Human beings cannot live without some kind of weather. Nature nudges. The common law is a regulatory system, and it will nudge, even if it allows people to have a great deal of flexibility.

21. In this light, choice architecture is inevitable. Human beings (or dogs or cats or horses) cannot wish it away. Any store has a design; some products are seen first, and others are not. Any menu places options at various locations. Television stations are placed on different positions on the dial, and strikingly, position matters, even when the costs of switching are vanishingly low; people tend to choose the station at the lower position.⁴⁰ A website has a design, which will affect what and whether people will choose.⁴¹
22. Nor can the state avoid nudging. Any government, even one that is or purports to be firmly committed to laissez-faire, has to establish a set of prohibitions and permissions, including a set of default entitlements, establishing who has what before bargaining begins. Recall that the rules of contract (as well as property and tort) provide a form of choice architecture for social ordering. It is true that choice architecture can maintain freedom of choice; it is also true that choice architects can at least aspire to neutrality. But choice architecture itself is inevitable, which means that it is pointless to object to it on ethical groups.

C. Spontaneous Orders and Visible Hands

23. At the same time, we can imagine the following view: Choice architecture is unavoidable, to be sure, but it is important if it is the product of nature or some kind of spontaneous order, rather than of conscious design, or of the action of any designer. Perhaps the law can build on that order; perhaps that law of contract, property, and tort do exactly that. Invisible-hand mechanisms⁴² often produce choice architecture. On a time-honored view, much of law is in fact

³⁸ Available at <http://moreintelligentlife.com/story/david-foster-wallace-in-his-own-words>

³⁹ Meghan R. Busse et al., Projection Bias in the Car and Housing Markets (2014), available at <http://www.nber.org/papers/w18212>

⁴⁰ See Gregory Martin and Ali Yurukoglu, Bias in Cable News: Real Effects and Polarization (2014), available at <http://www.nber.org/papers/w20798>

⁴¹ See Steve Krug, Don't Make Me Think Revisited: A Common Sense Approach to Web and Mobile Usability (2014).

⁴² For a superb discussion, see Edna Ullmann-Margalit, The Invisible Hand and the Cunning of Reason, 64 Social Research 181 (1997).

“customary law.” It codifies people’s actual practices, and it does not reflect any kind of dictation by public authorities.

24. Consider Hayek’s celebration of the “empiricist, evolutionary tradition,” for which “the value of freedom consists mainly in the opportunity it provides for the growth of the undesigned, and the beneficial functioning of a free society rests largely on the existence of such freely grown institutions. There probably never has existed a genuine belief in freedom, and there certainly been no successful attempt to operate a free society, without a genuine reverence for grown institutions, for customs and habits”⁴³
25. On this view, there is special reason, from the standpoint of freedom, in valuing forms of choice architecture that reflect the work of “grown institutions,” rather than designed ones. We might be comfortable with any nudging that reflects “customs and habits” but suspicious of any nudging that displays no reverence for them. Here, then, is a foundation for skepticism about any kind of social engineering; the skepticism might be applied to nudges as well as to mandates and bans.
26. Alternatively, some aspects of choice architecture might be the product of a genuinely random process, and a choice architect might intentionally opt for randomness, on the ground that it has a kind of neutrality.⁴⁴ We might be particularly concerned about governmentally-designed choice architecture, on the ground that public officials lack the information or the incentives to be trusted.
27. Even if the law of contract, property, and tort constitute forms of choice architecture, and even if it is not quite customary (and involves a degree of dictation and design), the relevant architecture can be made as flexible as possible and maintain a great deal of room for private ordering – and thus for freedom. To summarize a lengthy argument⁴⁵: the state, and the law, can provide the background rules for private interaction and decline to specify outcomes. Even if those rules turn out to nudge (as in the case of default rules), they are very different from social planning – on one view, far more modest and less dangerous.
28. To be sure, the criminal law will include some dictation; we will not deal with murder, assault, and rape with mere nudges. But perhaps the criminal law can restrict itself to prohibitions on force and fraud (and also play a role in correcting the standard market failures). A nation can certainly minimize the number of activities that it criminalizes. It might

⁴³ Hayek, *supra* note, at 234.

⁴⁴ See Jon Elster, *Solomonic Judgments* (1989).

⁴⁵ See Friedrich Hayek, *The Constitution of Liberty* (1976).

adopt a narrow account of the scope of the criminal law, perhaps focused on harm to others. And it might restrict any nudging to initiatives that operate in the service of the criminal law, narrowly conceived.

29. On certain assumptions, self-conscious choice architecture by the state is especially worrisome, because it is explicitly directed at achieving certain goals. But it is necessary to ask: What are those assumptions, and are they likely to be correct? Why and exactly when would spontaneous order be benign? (Is there some kind of social Darwinism here⁴⁶?) We might be able to agree that action by government poses dangers and risks while insisting that efforts to defend spontaneous orders and invisible hands run into well-understood problems and objections.⁴⁷
30. For example, a government that forbids racial discrimination, or that takes steps to ensure reasonable accommodation of disabled people, is not relying on an invisible hand. In any case, the argument for spontaneous orders seeks to restrict, above all, the coercive power of the state, not nudges as such. Whatever our theory of the legitimate domain of government, the most serious harms tend to come from mandates and bans (from genuine coercion), and not from nudges, which maintain freedom of choice.
31. What is so good about randomness? We should agree that a malevolent choice architect, aware of the power of nudges, could produce a great deal of harm. But by definition, random processes have the vice of arbitrariness.
32. It is true that spontaneous orders, invisible hands, and randomness can avoid some of the serious dangers, and some of the distinctive biases, that come from self-conscious nudging on the part of government.⁴⁸ If we are especially fearful of official mistakes – coming from incompetence or bad motivations – we will want to minimize the occasions for nudging.⁴⁹ And if we believe that invisible hand mechanisms promote welfare or freedom, we will not want to disturb their products, even if those products include nudges.
33. In my view, the strong position in favor of spontaneous orders and invisible hands cannot, in the end, be defended. But my goal here is not

⁴⁶ See *id.*

⁴⁷ Ullmann-Margalit, *supra* note.

⁴⁸ See Edward Glaeser, *Paternalism and Psychology*, 73 U. Chi. L. Rev. 133, 136-39 (2006).

⁴⁹ *Id.*

to justify that conclusion. The minimal point is that a degree of official nudging cannot be avoided. If we are committed to spontaneous orders and invisible hands, we will be committed to a specific role for government, one that will include a specified choice architecture and specified nudges.

III. The Diversity of Nudges and the Trap of Abstraction

34. To come to terms with the ethical questions, it is exceedingly important to bring first principles in contact with concrete practices. For purposes of orientation, it will be useful to give a more detailed accounting of potential nudges that might alter choice architecture.⁵⁰ One reason is to avoid *the trap of abstraction*, which is not simple to avoid in this setting, and which can create serious confusion when we are thinking about regulatory (or other) policy.

A. Motivating Nudges

35. The most obvious nudges consist of default rules, which establish what happens if people do nothing at all.⁵¹ Others include simplification (for example, of applications for job training or financial aid); disclosure of factual information (for example, calorie labels); warnings, graphic or otherwise (for example, on cigarette packages); reminders (for example, of bills that are about to become due); increases in ease and convenience (for example, through website design); uses of social norms (for example, disclosure of how one's energy use compares to that of one's neighbors); nonmonetary rewards, such as public recognition; active choosing (as in the question: *what retirement plan do you want?* or *do you want to become an organ donor?*); and precommitment strategies⁵² (through which people agree, in advance, to a particular course of conduct, such as a smoking cessation program).
36. It is important to acknowledge that some nudges preserve freedom of choice for a relevant population, while mandating action from some other population. Suppose, for example, that the government requires large employers to adopt automatic enrollment plans either for retirement or for health insurance.⁵³ If some, employees are nudged, but employers are coerced. Or suppose that the government requires chain restaurants or movie theaters to display calories to consumers.⁵⁴ If so,

⁵⁰ See Cass R. Sunstein, *Nudging: A Very Short Guide*, J Consumer Policy (2014).

⁵¹ Eric Johnson et al., *Defaults, Framing and Privacy: Why Opting In-Opting Out*, 13 Marketing Letters 5 (2002)

⁵² See Ian Ayres, *Carrots and Sticks* (2011).

⁵³ The Affordable Care Act in fact does this.

⁵⁴ The Affordable Care Act does this as well.

customers are nudged, but restaurants are coerced. Some nudges from government take the form of requiring some group X to nudge some group Y.

37. In behavioral science, it has become standard to distinguish between two families of cognitive operations: System 1, which is fast, automatic, and intuitive, and System 2, which is slow, calculative, and deliberative.⁵⁵ System 2 can and does err, but System 1 is distinctly associated with identifiable behavioral biases.⁵⁶ To be sure, there is, in some circles, intense controversy about the appropriate evaluation of the automatic system and about the extent to which it should be associated with error. Perhaps our intuitions usually work well in the situations in which we ordinarily find ourselves.⁵⁷ But there is no question that our intuitions often misfire, and that a good nudge could provide a great deal of help.
38. Some nudges, imposed by regulatory agencies, attempt to strengthen the hand of System 2 by improving the role of deliberation and people's considered judgments – as, for example, through disclosure of relevant information, debiasing,⁵⁸ and the use of precommitment strategies. Other nudges are designed to appeal to, or to activate, System 1 – as in the cases of graphic warnings.⁵⁹ Some nudges do not appeal to System 1, but work because of its operation – as, for example, where default rules have large effects because of the power of inertia.⁶⁰

⁵⁵ See Daniel Kahneman, *Thinking, Fast and Slow* (2011).

⁵⁶ See *id.*

⁵⁷ This position is vigorously defended in Gerd Gigerenzer et al., *Simple Heuristics That Make Us Smart* (2000); for general discussion, see Mark Kelman, *The Heuristics Debate* (2011). In my view, the outcome of this debate of this occasionally heated (and somewhat theological) debate does not have strong implications for policy, practice, or ethics. Everyone should agree that heuristics generally work well; that is why they exist. Everyone should also agree that in important cases, boundedly rational people make mistakes. When they make mistakes, some kind of nudge might help. To be sure, the best nudge may or may not involve education. See *infra*.

⁵⁸ See Christine Jolls and Cass R. Sunstein, *Debiasing Through Law*, 35 *J. Legal Stud.* 199 (2006)

⁵⁹ *R.J. Reynolds Tobacco Co. v. U.S. Food & Drug Admin.*, 823 F. Supp. 2d 36 (D.D.C. 2011), *rev'd*, 696 F.3d 1205 (D.C. Cir. 2012).

http://www.gpo.gov/fdsys/pkg/USCOURTS-dcd-1_11-cv-01482/pdf/USCOURTS-dcd-1_11-cv-01482-0.pdf

⁶⁰ See Johnson et al., *supra* note.

39. A nudge might be justified on the ground that it helps counteract a behavioral bias,⁶¹ and (as we shall see) some people object to such efforts, especially if they seem to target or to exploit System 1. But (and this is an important point) a behavioral bias is *not* a necessary justification for a nudge,⁶² and nudges need not target or exploit System 1 in any way. Disclosure of information can be helpful even in the absence of any bias. A default rule simplifies life and might therefore be desirable whether or not a behavioral bias is involved. A GPS is useful even for people who do not suffer from any such bias.⁶³
40. As the GPS example suggests, many nudges have the goal of *increasing navigability* – of making it easier for people to get to their preferred destination. Such nudges stem from an understanding that life can be simple or hard to navigate, and a goal of helpful choice architecture is desirable as a way of promoting simple navigation. To date, there has been far too little attention to the close relationship between navigability and (good) nudges. Insofar as the goal is to promote navigability, the ethical objections are greatly weakened and might well dissipate.
41. Nudges can have a substantial effect on both individual lives and social welfare. In Denmark, automatic enrollment in retirement plans has had a much larger effect than substantial tax incentives.⁶⁴ In the United States, efforts to inform consumers of how their energy use compares to that of their neighbors has had the same (significant) effect has a significant spike in the short-term cost of electricity.⁶⁵ Simplification of the financial aid form, to assist people who seek to attend college, has been found to have as large an effect, in promoting college attendance, as a several thousand dollar increase in financial aid.⁶⁶

⁶¹ See Christine Jolls and Cass R. Sunstein, *Debiasing Through Law*, 35 *J Legal Stud* 199 (2006).

⁶² This conclusion means that the highly illuminating discussion in RICCARDO REBONATO, *TAKING LIBERTIES* (2012), offers a mistaken definition.

⁶³ Conly, *supra* note, writes that for those who endorse nudging, the “assumption is that because our decision-making ability is limited, we need to use nonrational means to seduce people into doing what is good for them, and are trying to get people to act through the use of nonrational means.” *Id.* at 30. This is not the assumption that lies behind nudging, though perhaps some nudges can be understood in this way. See below.

⁶⁴ Raj Chetty et al., *Active vs. Passive Decisions and Crowdout in Retirement Savings Accounts: Evidence from Denmark* 38 (Nat’l Bureau of Econ. Research, Working Paper No. 18565, 2012), available at <http://www.nber.org/papers/w18565>

⁶⁵ See Hunt Alcott, *Social Norms and Energy Conservation*, 85 *J Public Econ* 1082 (2011).

⁶⁶ See Eric Bettinger et al., *The Role of Simplification and Information in College Decisions* (2009), available at <http://www.nber.org/papers/w15361>

B. Three Distinctions

42. It follows that for purposes of evaluating the ethical questions, three distinctions are particularly important. First, *paternalistic nudges* should be distinguished from what we might call *market failure nudges*. Some of the most familiar nudges are designed to protect people from their own mistakes; others are designed to respond to some kind of market failure.
43. In the latter category, we can identify *externality-reducing nudges*, *prisoner's dilemma nudges*, and *coordination*⁶⁷ *nudges*. If third parties are adversely affected, and if the choice architect's goal is to reduce those adverse effects, we should not speaking of using nudges to steer people so as to increase the likelihood that their decisions will make their own lives go better (by their lights). We should speak instead of reducing adverse third-party effects.
44. For nudges that fall in this category, the governing question should be: Do they increase social welfare, rightly understood⁶⁸? Cost-benefit analysis is the best available way of operationalizing that question, though it has significant gaps and limitations,⁶⁹ and though distributional considerations might turn out to be relevant.⁷⁰
45. As an ethical matter, market failure nudges should not be especially controversial in principle, though we might well worry over questions of effectiveness.⁷¹ In the face of a standard market failure, a mere nudge is usually not enough; coercion might well be justified (perhaps in the form of a corrective tax, perhaps in the form of a regulatory mandate). But a nudge might prove to be complementary to coercion, and in some ways, it might be a substitute.⁷²

⁶⁷ On coordination, with implications for productive nudges, see Edna Ullmann-Margalit, *Coordination Norms and Social Choice*, 11 *Erkenntnis* 143 (1977).

⁶⁸ Of course there is a great deal of dispute about how social welfare is rightly understood. For a valuable discussion, see Matthew Adler, *Welfare and Fair Distribution* (2011).

⁶⁹ See Cass R. Sunstein, *Valuing Life* (2014). Within economics, a pervasive question is whether interpersonal comparisons of utility are possible. See *Interpersonal Comparisons of Utility* (Jon Elster ed. 1991).

⁷⁰ See Cass R. Sunstein and Lucia Reisch, *Automatically Green*, 38 *Harv Env L Rev* 128 (2014).

⁷¹ See *id.*; Ryan Bubb and Richard Pildes, *How Behavioral Economics Trims Its Sails and Why*, 127 *Harv. L. Rev.* 1593 (2014).

⁷² Cass R. Sunstein and Lucia Reisch, *Automatically Green*, 38 *Harv Env L Rev* 128 (2014).

46. Second, *educative nudges* should be distinguished from nudges that *lack educative features*. Educative nudges attempt to inform people, so that they can make better choices for themselves. Other nudges are meant to help people without increasing their knowledge or understanding; default rules have this characteristic. It should be clear that in important respects, educative nudges should not be especially controversial, though their benefits might not justify their costs,⁷³ and though they can also run into problems of effectiveness.⁷⁴
47. Third, *nudges that enlist or exploit behavioral biases* should be distinguished from nudges that do no such thing. We have seen that some nudges enlist or exploit System 1 whereas other nudges appeal to System 2.⁷⁵ Efforts to target, or to benefit from, behavioral biases tend to be more controversial than efforts to appeal to deliberative capacities.⁷⁶
48. It follows that the most controversial nudges are paternalistic, noneducative, and designed to enlist or exploit behavioral biases.

C. Illicit Reasons and Transparency

49. It must be acknowledged that choice architecture can be altered, and that new nudges can be introduced, for illicit reasons. Indeed many of the most powerful objections to nudges, and to changes in choice architecture, are based on a judgment that the underlying motivations are illicit.⁷⁷ With these points, there is no objection to nudges as such; the objection is to the grounds for the particular nudges.
50. For example, an imaginable default rule might skew the democratic process by saying that voters are presumed to vote for the incumbent politician, unless they specify otherwise. Such a rule would violate principles of neutrality that are implicit in democratic norms; it would be unacceptable for that reason. Alternatively, a warning might try to frighten people about the supposedly nefarious plans of members of a minority group. Social norms might be used to encourage people to buy unhealthy products. In extreme cases, private or public institutions might try to nudge people toward violence.

⁷³ See George Loewenstein et al., Disclosure: Psychology Changes Everything, 6 Annual Review of Economics 391 (2014).

⁷⁴ See Lauren Willis, The Financial Education Fallacy, 101 Am. Econ. Rev. 429 (2011).

⁷⁵ See Daniel Kahneman, Thinking, Fast and Slow (2011).

⁷⁶ Riccardo Rebonato, Taking Liberties (2011).

⁷⁷ See Glaeser, *supra* note; Rebonato, *supra* note.

51. It must also be acknowledged that the best choice architecture often calls for active choosing.⁷⁸ Sometimes the right approach is to *require* people to choose, so as to ensure that their will is actually expressed. Sometimes it is best to *prompt* choice, by asking people what they want, without imposing any requirement that they do so.⁷⁹ A prompt is emphatically a nudge, designed to get people to express their will, and it might be unaccompanied by any effort to steer people in a preferred direction – except in the direction of choosing.
52. Choice architecture should be transparent and subject to public scrutiny, certainly if public officials are responsible for it. At a minimum, this proposition means that when such officials institute some kind of reform, they should not hide it from the public. If officials alter a default rule so as to promote clean energy or conservation, they should disclose what they are doing. Self-government itself requires public scrutiny of nudges. Such scrutiny is an important *ex ante* safeguard against harmful nudges; it is also an important *ex post* corrective. Transparency and public scrutiny can reduce the likelihood of welfare-reducing choice architecture. Nations should also treat their citizens with respect, and public scrutiny shows a measure of respect at the same time that it reduces the risk that nudges will intrude on either autonomy or dignity.
53. There is a question whether transparency and public scrutiny are sufficient rather than merely necessary. The answer is that they are not sufficient. We could imagine forms of choice architecture that would be unacceptable even if they were fully transparent; consider (transparent) architecture designed to entrench inequality on the basis of sex. Here again, the problem is that the goals of the relevant nudge are illicit. As we shall see, it is also possible to imagine cases of manipulation, in which the goals are not illicit, but in which the fact of transparency might not be sufficient to justify a nudge. A transparent nudge, announced in advance but taking the form of subliminal advertising of one or another kind, would run into legitimate objections about manipulation.

IV. “As Judged By Themselves”

A. The Basic Standard

⁷⁸ See Cass R. Sunstein, *Deciding By Default*, 162 U Pa L Rev 1 (2013).

⁷⁹ Kessler and Roth, *supra* note.

54. Recall that there have been recurrent ethical concerns about nudging, taken as such,⁸⁰ and that the principal concerns involve autonomy, dignity⁸¹ (sometimes described as “respect for persons”), manipulation,⁸² and learning.⁸³ In their most ambitious forms, the concerns lead to a conclusion that nudges can or do violate individual rights. Obviously this might be true if their goal is illicit, but it is correct to say that some kind of violation (for example, an insult to autonomy) can occur even without an illicit goal.⁸⁴
55. When third parties are not at risk, and when the welfare of choosers is all that is involved, the objective of nudging is to “influence choices in a way that will make choosers better off, *as judged by themselves*”⁸⁵ (italics in original). In many cases, that standard is straightforward to apply. If a GPS steers people toward a destination that is not their own, it is not working well. And if it offers them a longer and less convenient route, it will not make choosers better off by their own lights.
56. Many nudges can be understood in precisely the same terms; consider a reminder, a warning, or disclosure of relevant information. To enlist the “as judged by themselves” standard, we would have to take each nudge on its own. But the standard will often provide sufficient guidance.

B. Questions and Doubts

57. At the same time, it must be acknowledged that the standard raises normative, conceptual, and empirical challenges.

1. Objectively good lives?

⁸⁰ Edward L. Glaeser, *Paternalism and Psychology*, 73 U. Chi. L. Rev. 133, 136-39 (2006) (offering examples of how individuals’ beliefs and opinions can be manipulated); Joshua D. Wright & Douglas H. Ginsburg, *Behavioral Law and Economics: Its Origins, Fatal Flaws, and Implications for Liberty*, 106 Nw. U. L. Rev. 1033 (2012).

⁸¹ See Jeremy Waldron, *It’s All For Your Own Good*, New York Review of Books (2014), available at <http://www.nybooks.com/articles/archives/2014/oct/09/cass-sunstein-its-all-your-own-good/> Consider in particular this question: “Deeper even than this is a prickly concern about dignity. What becomes of the self-respect we invest in our own willed actions, flawed and misguided though they often are, when so many of our choices are manipulated to promote what someone else sees (perhaps rightly) as our best interest?”

⁸² See White, *supra* note.

⁸³ See Rebonato, *supra* note.

⁸⁴ See *id.*

⁸⁵ Thaler and Sunstein, *supra* note, at 5.

58. Some people believe that human lives can be objectively good or objectively bad, and that choosers can and do make objective mistakes about what makes their lives good. “Perfectionist”⁸⁶ approaches emphasize the importance and legitimacy of approaches to government that focus on increasing the likelihood that people will have good lives. (Of course there are many different forms of perfectionism, and I am bracketing many complexities here.)
59. For people who have this belief, the “as judged by themselves” standard is based on a fundamental mistake, which is that it allows the judgments of choosers to prevail even if they are wrong. Imagine, for example, that a chooser makes decisions that ensure a life that is short and unhealthy, or that is without either meaning or pleasure,⁸⁷ or that involves a great deal of suffering. It might be asked: Why should choice architects defer to choosers in such circumstances?
60. This question raises serious questions within political philosophy, which I cannot answer here.⁸⁸ To the extent that choice architects defer to choosers, it might be because of a (moral) judgment that choosers have ultimate sovereignty over their own lives, or it might be because of their own humility – their understanding that they have epistemic disadvantages as compared with those whose own lives are at stake. For present purposes, the central point is that insofar as nudgers adopt the “as judged by themselves” standard, they reject perfectionism, and they do so on principle.

2. Ex ante or ex post?

61. The “as judged by themselves” standard raises this question: Do we ask about choosers’ judgments before the nudge, or instead after? Choosers’ ex ante judgments might diverge from their ex post judgments. If choosers’ judgments are constructed by the nudge, then *choice architects might be engineering the very judgment from which they are claiming authority*. That is a serious problem for the “as judged by themselves” standard.
62. Suppose, for example, that with a “green” default rule – one that provides an environmentally friendly energy provider, subject to opt out

⁸⁶ A form of liberal perfectionism is defended in Joseph Raz, *The Morality of Freedom* (1985).

⁸⁷ On the importance of purpose and pleasure, see Paul Dolan, *Happiness By Design* (2014).

⁸⁸ Relevant discussion can be found in Adler, *supra* note; Amartya Sen, *Commodities and Capabilities* (1999); Martha Nussbaum, *Creating Capabilities* (2013). For a short, vivid set of objections to perfectionism, see Conly, *supra* note.

in favor of a cheaper but environmentally inferior provider – choosers are perfectly content. But suppose that they would also be content with the opposite default rule. Which judgments matter? Wherever the nudge influences choosers’ judgments, that question raises serious puzzles.

63. In most cases, the nudge is unlikely to affect choosers’ judgments; they will be the same *ex ante* and *ex post*. But when they differ, the standard does become more difficult to apply. One option would be to use active choosing to see what people actually want. Another would be to explore the number of opt-outs under different default rules.⁸⁹ A third would be to attempt a more direct inquiry into people’s welfare under different forms of choice architecture, though admittedly any such inquiry raises challenges of its own.⁹⁰

3. Preferences about preferences

64. An additional question is raised by the fact that people do not only have preferences (or first-order preferences); they also have preferences about their preferences (or second-order preferences).⁹¹ People might want to eat delicious but fattening foods, or spend monthly salary every month, but they might want not to want those things. In applying the “as judged by themselves standard,” should choice architects consult first-order or second-order preferences?
65. Some imaginable cases are difficult, but in general, the answer is straightforward: If second-order preferences reflect System 2 – understood, in this context, as people’s reflective judgments as opposed to their impulses – there is a strong argument that they have authority.

4. Informed judgments and empirical puzzles

66. When we ask about choosers’ judgments, what kind of information do we expect choosers to have? It makes sense to say that choice architects should defer to choosers’ informed judgments, rather than their uninformed ones. But if choice architects are loosened from choosers’ actual judgments, and asking what choosers would do if they were informed, there is a risk that choice architects will be relying on their own values and beliefs, rather than choosers’ own.
67. In any case, such architects might lack sufficient information to know whether informed choosers deem themselves to be better off. It might

⁸⁹ Thaler and Sunstein, *supra* note.

⁹⁰ See Dolan, *supra* note.

⁹¹ See Harry Frankfurt, *Freedom of the Will and the Concept of a Person*, 68 *J Phil* 5 (1971).

not be at all simple for outsiders to compare (from choosers' point of view) the various outcomes that stem from different nudges.

68. In some cases, these points might raise serious conceptual and empirical challenges. Nonetheless, the idea of choosers' informed judgments serves as the lodestar, and it imposes real discipline.⁹² Certainly choice architects should be focused on the welfare of choosers, rather than their own. (In a well-functioning market system, that focus is essentially guaranteed for market participants, at least under optimistic assumptions.⁹³)

6. Self-control

69. There are also hard questions about how to handle the “as judged by themselves” standard in the face of self-control problems. Suppose that someone faces such problems and is aware of that fact – but nonetheless wishes, at Time 1, to give into his impulses. Do we look to the assessment of (1) the alcoholic, who wants that beer, (2) the would-be former alcoholic, who wants to quit, or (3) the actual former alcoholic, who is grateful to have been nudged away from alcoholism? In some ways, this question replicates those involve ex ante vs. ex post judgments, and also those involving preferences about preferences. But insofar as the focus is on self-control problems, the issue is distinctive.
70. It is reasonable to emphasize that no former alcoholic regrets the “former.” For that reason, there is a strong argument that the “as judged by themselves” criterion should be taken to refer to the judgment of the person who is no longer in the grip of an addiction. Nonetheless, there can be a thin line between a self-control problem and a legitimate focus on short-term pleasure; the question deserves more extended treatment. No choice architect should engage in a program of nudging that disregards the important of short-term pleasures, or pleasures in general, which are of course crucial parts of good lives.⁹⁴

C. Bounded Rationality: Education and its Discontents

⁹² As noted, difficult questions might be raised where subjective well-being departs from objective well-being; I am bracketing those questions. See Martha Nussbaum, *Creating Capabilities* (2013); Amartya Sen, *Development As Freedom* (2000). Note also that in the presence of “affective forecasting errors,” an emphasis on subjective well-being leads to a focus on what, in fact, makes people (subjectively) better off, not on what they anticipate will make them better off.

⁹³ Optimistic, not realistic. See Oren Bar-Gill, *Seduction By Contract* (2011).

⁹⁴ See Dolan, *supra* note.

71. Decades of work in behavioral science have specified how human beings depart from full rationality.⁹⁵ It would not be helpful or correct to say that people are “irrational.” It is more helpful and correct to say that they are human and to draw attention to “bounded rationality” (and to specify what they means) Most obviously, people often lack important information. They are also subject to specific biases. For example, most people tend to be unrealistically optimistic.⁹⁶ People also show “present bias,” focusing on the short-term and downplaying the future.⁹⁷ People do not deal well with probability, in part because they use heuristics, or mental shortcuts, that sometimes lead them in unfortunate directions.⁹⁸ With respect to probability, people’s intuitions can go badly wrong, in the sense that they can produce serious mistakes, including life-altering ones.⁹⁹
72. It is true, of course, that in the face of error, education well might be the best response.¹⁰⁰ Much of the time, the first and the best line of defense is education, which might itself be characterized as a nudge, and which certain counts as a form of choice architecture. We have seen that educative nudges are an important part of the repertoire of the choice architect. Thus Jeremy Waldron writes: “I wish, though, that I could be made a better chooser rather than having someone on high take advantage (even for my own benefit) of my current thoughtlessness and my shabby intuitions.”¹⁰¹
73. But education has its limits. People benefit from default rules with respect to cell phones, tablets, health insurance policies, and rental car agreements. (To be sure, not all such rules are beneficial.) If people had to obtain sufficient education on all of the underlying issues, they would

⁹⁵ See Daniel Kahneman, *Thinking Fast and Slow* (2011).

⁹⁶ See Tali Sharot, *The Optimistic Bias* (2011).

⁹⁷ For references and discussion, see Sunstein, *Why Nudge?*, supra note.

⁹⁸ See Kahneman, supra note.

⁹⁹ For a powerful demonstration, see Daniel Chen et al., *Decision-Making under the Gambler’s Fallacy: Evidence from Asylum Judges, Loan Officers, and Baseball Umpires* (2014), available at <http://users.nber.org/~dlchen/papers/ColdHand.pdf>

¹⁰⁰ See Gerd Gigerenzer, *Risk Savvy* (2014). It is certainly reasonable to think that education or educative nudges have large advantages over (for example) default rules, because the former provide people with a kind of “stock” from which they can make their own decisions. But in some cases, default rules are preferable because they do not impose significant demands on choosers, because they are more effective, and because the area is one for which education is not particularly important. For discussion, see Cass R. Sunstein, *Choosing Not To Choose* (2015).

¹⁰¹ See Jeremy Waldron, *It’s All For Your Own Good*, NY Rev of Books (2014), available at <http://www.nybooks.com/articles/archives/2014/oct/09/cass-sunstein-its-all-your-own-good/>

quickly run out of time. In many cases, a default rule is desirable, because it would preserve desirable outcomes (again, from the standpoint of choosers themselves) without requiring people to take the functional equivalent of a course in (say) statistics or finance.¹⁰² There is a recurring question whether in particular circumstances, the costs of education justify the benefits. For those who are engaged in many activities (including shopping), it would be impossibly demanding to insist on the kind of education that would allow active choices about all relevant features. Default rules may well be best. Everything depends on the facts, but there is a good argument that with respect to certain retirement issues, default rules are preferable to financial education.¹⁰³

74. Though choice architecture and nudging are inevitable, some particular nudges are certainly avoidable.¹⁰⁴ A government might decide not to embark on a campaign to discourage smoking or unhealthy eating. It could ignore the problem of obesity. It could refrain from nudging people toward certain investment behavior. To that extent, it is reasonable to wonder whether government should minimize nudging.¹⁰⁵ If we distrust the motives of public officials, or believe that their judgments are likely to go wrong, we will favor such minimization.¹⁰⁶

V. Seven Objections

75. Here is a more detailed catalogue of potential objections to nudges as such.
- (a) Some nudges are paternalistic; in a free society, respectful of freedom of choice and self-determination, there should be a strong presumption against paternalism.
 - (b) Some nudges intrude on people's autonomy and are unacceptable for that reason.
 - (c) Some nudges turn out to be coercive, even if they preserve freedom of choice as a formal or technical matter.
 - (d) Some nudges insult people's dignity; they are infantilizing; they treat people as children.¹⁰⁷ The idea of the "nanny state" captures this objection.

¹⁰² See Lauren Willis, *The Financial Education Fallacy*, 101 *Am Econ Rev* 429 (2011).

¹⁰³ See Willis, *supra* note.

¹⁰⁴ See Glaeser, *supra* note.

¹⁰⁵ See *id.*

¹⁰⁶ *Id.*

¹⁰⁷ See Joel Feinberg, *Legal Paternalism*, in *Paternalism* 3 (R. Darforius ed. 1983): "If adults are treated as children they will come in time to be like children. Deprived of the right to choose for themselves, they will soon lose the power of rational judgment and decision." This is a claim about deprivation of the right to choose, but it could be adapted to apply to default rules as well. See Jeremy Waldron, *It's All For Your Own Good*, NY

(e) Some nudges could count as forms of manipulation.¹⁰⁸ It is relevant in this regard that nudging might operate without triggering people’s conscious awareness and is not sufficiently transparent. Consider, for example, “negative option marketing,” by which people who purchase certain products find themselves enrolled in programs for which they pay a monthly fee.¹⁰⁹ Nudges of this kind might be said to operate “behind people’s backs.” (The concept of manipulation is not self-defining and deserves considerable attention.¹¹⁰)

(f) Some nudges impede or at least do not promote learning. Educative nudges have the advantage of helping people to become better choosers. But as we have seen, some nudges lack that advantage, and they might discourage people from learning their own.

(g) Choice architects may err, especially when they work for government, and for that reason, it is best to avoid nudging (to the extent that this is possible). We should not trust choice architects who are on the public payroll, because they may now know what they are doing, and because they might well be focused on their own interests, rather than ours.

76. It is important not to take these concerns as all-purpose objections to efforts to improve choice architecture. Does any of these objections make sense as applied to initiatives designed to promote active choosing? To inform consumers of the caloric content of food, to remind people that a bill is due, or to ask people whether they want to enroll in a retirement plan? But let us take the objections in sequence.

A. Paternalism

Rev of Books (2014), available at <http://www.nybooks.com/articles/archives/2014/oct/09/cass-sunstein-its-all-your-own-good/>. Thus Waldron writes: “What becomes of the self-respect we invest in our own willed actions, flawed and misguided though they often are, when so many of our choices are manipulated to promote what someone else sees (perhaps rightly) as our best interest? . . . Nudging takes advantage of my deficiencies in the way one indulges a child. The people doing this (up in Government House) are not exactly using me as a mere means in violation of some Kantian imperative. They are supposed to be doing it for my own good. Still, my choosing is being made a mere means to my ends by somebody else—and I think this is what the concern about dignity is all about.”

¹⁰⁸ As we shall see, manipulation is troublesome, on ethical grounds, if it runs into a foundational commitment of some kind – for example, to autonomy or dignity. I explore it separately because it raises distinctive considerations.

¹⁰⁹ See 16 C.F.R. § 425.1 (2012) (regulating the use of prenotification negative-option plans); FTC, Negative Options 2 (2009), available at <http://www.ftc.gov/os/2009/02/P064202negativeoptionreport.pdf> (describing four types of plans that could be classified as negative-option marketing).

¹¹⁰ See note supra.

77. Choice architecture may or may not be paternalistic. But it is true that nudges can be seen as a form of “libertarian paternalism” insofar as they attempt to use choice architecture to steer choosers in directions that will promote their welfare (again, as judged by choosers themselves).¹¹¹
78. Recall that this is a distinctive form of paternalism in the sense that it is (a) soft and (b) means-oriented.¹¹² It is soft insofar as it avoids coercion or material incentives, and thus fully maintains freedom of choice.¹¹³ It is means-oriented insofar as it does not attempt to question or alter people’s ends. Like a GPS, it respects those ends (subject to the various complexities discussed above). To those who object to paternalism, the most serious concerns arise in the face of coercion (where freedom of choice is blocked) and when social planners, or choice architects, do not respect people’s ends.¹¹⁴ To this extent, nudges aspire to avoid some of the standard ethical objections to paternalism.¹¹⁵
79. Nonetheless, some skeptics object to paternalism as such.¹¹⁶ Perhaps people are the best judges not only of their ends, but also of the best means to achieve those ends, given their own tastes and values. (People might reject the route suggested by the GPS on the ground that they prefer the scenic alternative; the GPS might not easily capture or serve their ends.) Moreover, the distinction between means and ends is not always simple and straightforward. One question is the level of abstraction at which we describe people’s ends. If we describe people’s ends at a level of great specificity – eating that brownie, having that cigarette, texting while driving – then people’s means effectively *are* their ends. The brownie is exactly what they want; it is not a means to anything at all (except the experience of eating it).
80. If, by contrast, we describe people’s ends at a level of high abstraction – “having a good life” – then nearly everything is a means to those ends. But if we do that, then we will not be capturing people’s actual concerns; we will be disregarding what matters to them. These points do raise some problems for those who favor a solely means-oriented form

¹¹¹ Thaler and Sunstein, *supra* note.

¹¹² See Cass R. Sunstein, *Why Nudge?* (2014).

¹¹³ On some of the complexities here, see *id.*

¹¹⁴ This is the fundamental concern in John Stuart Mill, *On Liberty* (2d ed. 1863), *reprinted in* *The Basic Writings of John Stuart Mill: On Liberty, The Subjection of Women, and Utilitarianism* 3, 11-12 (Dale E. Miller ed., 2002).

¹¹⁵ The debate over perfectionism is of course relevant here. See Conly, *supra* note.

¹¹⁶ Rebonato, *supra* note; Joshua D. Wright & Douglas H. Ginsburg, *Behavioral Law and Economics: Its Origins, Fatal Flaws, and Implications for Liberty*, 106 *Nw. U. L. Rev.* 1033 (2012).

of paternalism. They must be careful to ensure that they are not describing people's ends at a sufficiently high level of abstraction as to misconceive what people care about.¹¹⁷

81. But insofar as a GPS is a guiding analogy, it is not easy to see nudges as objectionably paternalistic. Many nudges are entirely focused on helping people to identify the best means for achieving their preferred ends. Consider cases in which people are *mistaken about facts* (with respect to the characteristics of, say, a consumer product or an investment). If a nudge informs them, then it is respecting their ends. Or suppose that certain product characteristics are in some sense shrouded, and the nudge helps people to see them for what they are. Or suppose that people suffer from a behavioral bias – perhaps because they use the availability heuristic, perhaps because of unrealistic optimism. A nudge that corrects their mistake can help them to achieve their ends.
82. To be sure, some behavioral biases are not easy to analyze in these terms. If people suffer from present bias, is a nudge a form of paternalism about means? Suppose that people eat high calorie food, or drink a great deal, or fail to exercise, because they value today and tomorrow, and not so much next year or next decade. If a nudge succeeds in getting people to focus on their long-term interests, it might increase aggregate (intrapersonal) welfare over time. But is such a nudge focused solely on means? If a person is seen a series of selves extending over time, the choice architect is effectively redistributing welfare from earlier selves to later ones (and by hypothesis maximizing welfare as well). But it is not clear that we can speak, in such cases, of means paternalism. And if a person is seen as continuous over time, and not a series of selves, efforts to counteract present bias are, by hypothesis, undermining the ends of the chooser at the time of choice.
83. Let us bracket the most difficult issues and acknowledge that some forms of choice architecture count as paternalistic. Is that a problem? One reason for regulators and other policymakers to reject paternalism involves welfare: Perhaps people are the best judges of what will promote their interests, and perhaps outsiders will blunder (as John Stuart Mill believed¹¹⁸). Consider Hayek's remarkable suggestion that "the awareness of our irremediable ignorance of most of what is known to somebody [who is a planner] is *the chief basis of the argument for liberty*."¹¹⁹ A form of paternalism that maintains freedom of choice, and

¹¹⁷ See Rebonato, *supra* note.

¹¹⁸ JOHN STUART MILL, ON LIBERTY 8 (Kathy Casey ed., 2002) (1859).

¹¹⁹ Friedrich Hayek, The Market and Other Orders, in *The Collected Works of F. A. Hayek* 384 (Bruce Caldwell ed., 2013).

that is focused on means, is less likely to be objectionable on welfare grounds, certainly if we attend to behavioral biases.

84. In fact it is possible that welfarists should ultimately embrace coercive paternalism, at least in the face of such biases.¹²⁰ When paternalism would improve welfare, welfarists should support paternalism. For welfarists, paternalism should be evaluated on a case-by-case basis – unless there is some systematic, or rule-welfarist, reason to support a principle or presumption against paternalism.¹²¹
85. Perhaps there is good reason for such a presumption, rooted in a judgment that choosers are likely to have better information than choice architects.¹²² But in some cases, that judgment is incorrect, because choosers lack knowledge of facts. Information-providing nudges are a natural corrective. In some cases, a good default rule – say, automatic enrollment in pension programs – is hard to reject on welfarist grounds. To be sure, active choosing might be better, but that conclusion is not obvious. Welfarists might well be inclined to favor choice-preserving approaches, on the theory that individuals usually well know what best fits their circumstances, but the fact that a default rule has a paternalistic dimension should not be decisive against it.
86. Another reason to reject paternalism involves autonomy and the idea of respect for persons. Stephen Darwell writes that the “objectionable character of paternalism of this sort is not that those who seek to benefit us against our wishes are likely to be wrong about what really benefits us. . . . It is, rather, primarily a failure of respect, a failure to recognize the authority that persons have to demand, within certain limits, that they be allowed to make their own choices for themselves.”¹²³ This brings us to the next objection.

B. Autonomy

87. Do nudges intrude on autonomy? Autonomy requires informed choices, and many nudges are specifically designed to ensure that choices are informed.¹²⁴ In the face of a behavioral bias, or some kind of systematic mistake (by the actor’s own reflective lights), it is hardly clear that a

¹²⁰ See Sarah Conly, *Against Autonomy* (2012).

¹²¹ See Cass R. Sunstein, *Nudges v. Shoves*, 127 Harv L Rev Forum 210 (2014).

¹²² See *id.*

¹²³ See Stephen Darwell, *The Value of Autonomy and the Autonomy of the Will*, 116 *Ethics* 263, 269 (2006).

¹²⁴ See George Loewenstein et al., *Disclosure: Psychology Changes Everything*, 6 *Annual Review Economics* 391 (2014).

nudge infringes on autonomy, rightly understood.¹²⁵ When they help correct some kind of bias, nudges might well promote people's autonomy. We might identify autonomy with people's reflective judgments, and many nudges operate in the interest of autonomy, so understood.

88. It is also important to see that autonomy does not require choices everywhere; it does not justify an insistence on active choosing in all contexts. There is a close relationship between time-management and autonomy. People should be allowed to devote their attention to the questions that, in their view, deserve that attention. If people have to make choices everywhere, their autonomy is reduced, if only because they cannot focus on those activities that seem to them most worthy of their attention.¹²⁶
89. It is nonetheless true that on grounds of autonomy (as well as welfare), the best choice architecture often calls for active choosing. Even though they preserve freedom of choice, default rules might intrude on autonomy, certainly if they do not track people's likely choices. The problem is that because of the force of inertia, people might not reject harmful defaults.¹²⁷ If so, there is arguably an intrusion on their autonomy, because they will end up with outcomes that they did not specifically select. Consider, for example, a default rule that says that if you do not indicate otherwise, you are presumed to be a member of the National Patriots Party, or to want your estate to go to the Vatican, or prefer your organs to go to people with the right skin color. Even though people can opt out, default rules can intrude on autonomy insofar as they impose that burden on people – and insofar as the particular rules a) might stick because of that very burden and b) do not reflect what informed people would like.
90. Whether the interest in autonomy calls for active choosing, as opposed to reliance on a default rule, depends on the circumstances. Along some dimensions, default rules are actually superior to active choosing on autonomy grounds. If people choose not to choose, or if they would make that choice if asked, it is an insult to their autonomy to force them to choose.¹²⁸ And if people would like to choose, a default rule does not

¹²⁵ See Conly, *supra* note, at 36: “Even if we accept that individuals have rights, and thus claims not to be harmed by others in certain ways, or to have (yet) others defend them in these claims, why would there be such a right here, where the point of the action is to help the person achieve what in the long run, he wants, and what he would want not if he were not a flawed thinker?”

¹²⁶ See Sendhil Mullainathan and Eldar Shafir, *Scarcity* (2013).

¹²⁷ See Rebonato, *supra* note.

¹²⁸ See Cass R. Sunstein, *Choosing Not to Choose*, 64 *Duke LJ* 1 (2014).

deprive them of that choice; they can reject the default. Even in the face of inertia, many people will do so.¹²⁹

91. Preservation of freedom of choice is not sufficient, but it goes some distance toward ensuring that people's autonomy is respected. So does a requirement that any paternalistic nudges focus on people's own ends and otherwise have legitimate goals. But with respect to autonomy, a continuing problem lies in the possibility of manipulation; I will turn to that problem below.

C. Coercion

92. If choice architects coerce people, they are no longer merely nudging.¹³⁰ But skeptics might again emphasize that with some nudges, the problem of coercion can be avoided. We have seen that because of the power of inertia, people might accept (passively) a default rule even though they have no enthusiasm for the outcome that it produces, and would reject that outcome if they focused on the issue involved.¹³¹
93. We should doubt whether such situations are properly described as involving coercion. No one is being forced to do anything. But there is certainly a risk that a default rule will produce harmful results even though people have not affirmatively consented to the actions that led to them. Choice architects need to take account of that risk. But so long as freedom of choice is maintained and real, coercion is not involved.

D. Dignity

94. The idea of "dignity" is complex and contested. We might begin by suggesting that the antonym of coercion is freedom; the antonym of dignity is humiliation.¹³² Some nudges might seem to compromise dignity and respect for persons. As we shall see, this objection is both interesting and important, especially when it is combined with a concern about manipulation.¹³³ Imaginable forms of choice architecture could indeed undermine dignity.

¹²⁹ See *id.*

¹³⁰ Recall that an initiative might require its objects (employers, advertisers) to nudge others (employees, consumers); consider a mandatory fuel economy label.

¹³¹ For an excellent discussion, see Rebonato, *supra* note.

¹³² See Avishai Margalit, *The Decent Society* (1998).

¹³³ See Waldron, *supra* note.

95. There are of course large questions about the place of dignity in ethics and about the appropriate specification of the basic idea.¹³⁴ On one (admittedly unconventional) view, dignity is properly part of an assessment of welfare. If people feel humiliated, or feel that they have been treated disrespectfully, they suffer a welfare loss. That loss might be extremely serious. In any assessment of welfare consequences, such a loss must be considered. It might turn out to be exceedingly important – and to argue against particular nudges.
96. A good welfarist should also acknowledge that an offense to dignity is qualitatively distinct¹³⁵; in its nature, it is a different kind of loss from the loss of (say) money, or an opportunity to visit a beach. But on the welfarist view, a dignity loss is just one kind of loss, to be weighed against the other goods that are at stake. Suppose, for purposes of argument,¹³⁶ that a graphic and highly emotional appeal, triggering strong emotions (System 1) in order to discourage people from smoking, is plausibly seen as an offense to dignity – as a way of treating smokers disrespectfully (and perhaps infantilizing them). Some smokers might so regard such an appeal and object for that reason. A welfarist might be willing to support the emotional appeal, notwithstanding the relevant loss, if it saves a significant number of lives.
97. On another view, an insult to dignity is not merely part of a welfarist calculus. Such an insult does not depend on people's subjective feelings, and it is a grave act, perhaps especially if it comes from government. An insult to dignity should not be permitted unless (perhaps) it has an overwhelmingly strong justification. If we endorse this view, it is especially important to ask whether nudges offend human dignity.
98. To return to my general plea: The force of the objection depends on the particular nudge. A GPS insults no one's dignity. Disclosure of factual information can hardly be seen as an offense to dignity – certainly if the information is useful and not based on a false and demeaning belief that people need it.

¹³⁴ See Margalit, *supra* note; Michael Rosen, *Dignity: Its History and Meaning* (2012); Charles Beitz, *Human Dignity in the Theory of Human Rights*, 41 *Phil & Pub. Affairs* 259 (2013); Thomas Christiano, *Two Conceptions of Human Dignity as Persons* (unpublished manuscript 2008).

¹³⁵ On qualitative distinctions and their importance, see Elizabeth Anderson, *Value in Ethics and Economics* (1993).

¹³⁶ For discussion, see Kyle Rozema, *Economic Theory Lost in Translation: Will Behavioral Economics Reshape the Compelled Commercial Speech Doctrine?* (2013), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2511604

99. But we can easily imagine nudges that would offend one or another conception of dignity. Consider a public health campaign, directed at the prevention of obesity, that stigmatized and humiliated people who are overweight, by portraying them in a demeaning light.¹³⁷ Or consider, as a somewhat more difficult case, an antismoking campaign that did the same for smokers. Here again, the fact that nudges preserve freedom of choice, and do not require anyone to do anything, should not be taken as a kind of license to do anything at all. It is possible to imagine public education campaigns that offend dignity, though admittedly the more familiar real-world campaigns do not have anything approaching that vice.
100. It might also count as an insult to dignity, and a form of infantilization, if the government constantly reminds people of things that they already know. Every child, and everyone who was once a child, can recall this form of infantilization, and it is not always absent from adult life as well. If people are informed of the same thing every hour or even every day (say, by their spouse, by their doctor, or by some public official), they might legitimately feel that their dignity is not being respected.
101. The same things can be said about reminders, warnings, and uses of social norms.¹³⁸ If, for example, choice architects refer to norms, to let people know what most people do, they are not likely to be humiliating anyone.¹³⁹ In some cases, however, the concern about dignity might become more serious. If people are constantly reminded that a due date is coming, they might feel as if they are being treated like children. Warnings can run into the same concern insofar as they are repetition or condescending, or (are meant to) trigger strong emotions instead of merely giving people a sense of factual realities.¹⁴⁰
102. Here as well, there is no objection to the relevant nudges in the abstract, but there is an objection to imaginable nudging. At the same time, it must be emphasized that the relevant offense to dignity – coming from unwelcome and numerous reminders – is relatively minor, and from the standpoint of the concerns that have produced the focus on dignity in the Western political tradition, it is laughably modest.¹⁴¹
103. What is the relationship between dignity and default rules? If an employer automatically enrolls employees into retirement and health

¹³⁷ I am grateful to Gertrude Lubbe-Wolff for this example.

¹³⁸ On the functions of norms, see Edna Ullmann-Margalit, *The Emergence of Norms* (1976).

¹³⁹ See Alcott, *supra* note.

¹⁴⁰ See *id.*

¹⁴¹ See note *supra*.

care plans, dignity is hardly at risk. If a cell phone company adopts a series of defaults for the phone and the contractual arrangement, nothing need be amiss in terms of dignity.

104. But we could imagine harder cases. Suppose that the government insisted on “default meals” in various restaurants, so that people would be given certain healthy choices unless they specifically chose otherwise. Put to one side the fact that with respect to restaurants, this approach is a mandate, not a mere nudge. The reasonable response is: Why shouldn’t a free people be asked to select what they want? Or suppose that a government specified a “default exercise plan” for adults, so that they would be presumed to want to engage in certain activities unless they opted out. People might offer the same reasonable response, perhaps with considerable agitation.
105. Note that default rules of this kind might be objectionable for both welfarists and nonwelfarists. Welfarists might want to focus on people’s subjective feelings. If people believe that they are being treated as children, and if they object to that treatment, would count in the assessment. Nonwelfarists would insist that the offense to dignity is objectionable even if it has some kind of welfarist justification.
106. In extreme situations, default rules could indeed be a serious affront to dignity. If so, there should be a strong presumption against them (whatever our foundational commitments).¹⁴² But it would be a mistake to use extreme situations, or (barely?) imaginable cases, as a reason to challenge default rules in general. People are not treated disrespectfully if an institution adopts a double-sided default for printing, or if they are automatically enrolled in health insurance or retirement plans. The objection from dignity has far more force in the abstract than in the context of all, or nearly all, real-world cases in which default rules are actually at work.

E. Manipulation

107. To deal with this objection, we need to say something about the complex idea of “manipulation.”¹⁴³ It should be clear that an action does not count as manipulative merely because it is an effort to alter people’s behavior. If you warn a driver that he is about to get into a crash, you are not engaged in manipulation. The same is true if you remind

¹⁴² Perhaps the presumption could be rebutted with a sufficiently strong consequentialist justification – as, for example, by showing that many lives would be saved with the appropriate default.

¹⁴³ For helpful discussion, see White, *supra* note.

someone that a bill is due. A calorie label and an energy efficiency label are not ordinarily counted as forms of manipulation.¹⁴⁴

108. The idea of “manipulation” can be applied to many kinds of behavior; it is not clear that it is a unitary concept, or that we can identify necessary and sufficient conditions.¹⁴⁵ It seems clear that manipulation takes various forms, even if they are only loosely identified with one another.¹⁴⁶ An organizing idea is that when one is being manipulated, one is being treated as a kind of “puppet on a string.”¹⁴⁷ No one wants to be someone’s puppet, and it is especially bad to be a puppet of government. But T. M. Wilkinson is correct to say that “one should not rush to judgement when trying to decide whether nudging manipulates.”¹⁴⁸

1. Manipulation and conscious choice

109. An action can be counted as *manipulative* if it attempts to influence people subconsciously or unconsciously, in a way that undermines their capacity for conscious choice. Consider some variations on this idea. On Wilkinson’s account, manipulation “is a kind of influence that bypasses or subverts the target’s rational capacities.”¹⁴⁹ Wilkinson urges that manipulation “subverts and insults a person’s autonomous decision making,” in a way that treats its objects as “tools and fools.”¹⁵⁰ He

¹⁴⁴ A qualification is necessary. If a disclosure requirement focuses on one of many aspects of a situation, and fixes people’s attention on that aspect, a charge of manipulation would not be unreasonable. Consider the controversy over the idea that sellers should have to disclose that food has genetically modified organisms (GMOs). See Charles Noussair et al., DO CONSUMERS REALLY REFUSE TO BUY GENETICALLY MODIFIED FOOD?, 114 *Economic Journal* 102 (2004). For those who object to compulsory labeling about GMOs, there is a plausible claim that labels are a form of manipulation, activating public concern where there is no objective reason for that concern.

¹⁴⁵ For a number of valuable treatments, see *Manipulation* (Christian Coons and Michael Webster eds. 2014)

¹⁴⁶ A valuable cautionary note: “People can be manipulated when they go shopping, strike contracts, vote, study at school, visit their doctors, decide whether to have sex or take turns to do the housework. A full account would have to cope with the enormous variety of sites and methods of manipulation. Indeed, we do not have such an account.” *Id.* at 344.

¹⁴⁷ See Wilkinson, *supra* note, at 342.

¹⁴⁸ Wilkinson, *supra* note, at 342.

¹⁴⁹ Christian Coons and Michael Webster, Introduction, in *Manipulation* 11 (Christian Coons and Michael Webster eds. 2014)

¹⁵⁰ Wilkinson, *supra* note, at 145.

thinks that “manipulation is intentionally and successfully influencing someone using methods that pervert choice.”¹⁵¹

110. In a similar account, Ruth Faden and Tom Beauchamp define psychological manipulation as “any intentional act that successfully influences a person to belief or behavior by causing changes in mental processes *other than those involved in understanding*.”¹⁵² Joseph Raz suggests that “Manipulation, unlike coercion, does not interfere with a person’s options. Instead it perverts the way that person reaches decisions, forms preferences or adopts goals.”¹⁵³
111. Of course the idea of “perverting” choice, or people’s way of reaching decisions or forming preferences, is not self-defining; it is often taken to refer to methods that do not appeal to, or produce, conscious deliberation. If so, the objection to manipulation is that it “infringes upon the autonomy of the victim by subverting and insulting their decision-making powers.”¹⁵⁴
112. The objection certainly applies to lies, which attempt to alter behavior not by engaging people on the merits and asking them to decide accordingly, but by enlisting falsehoods in the service of the liar’s goals. A lie is disrespectful to its objects, not least if it attempts to exert influence without asking people to make a deliberate choice in light of relevant facts. In harder cases, the challenge is to concretize the ideas of “subverting” and “insulting.”
113. Subliminal advertising should be deemed manipulative and insulting, because it operates “behind the back” of the person involved, without appealing to his conscious awareness. People’s decisions are affected in a way that bypasses their own deliberative capacities. If this is the defining problem with subliminal advertising, we can understand why involuntary hypnosis would also count as manipulative. But almost no one favors subliminal advertising, and to say the least, the idea of involuntary hypnosis does not have much appeal. The question is whether admittedly taboo practices can shed light on interventions, including nudges, that might be able to command broader support.
114. Some forms of framing could plausibly be counted as manipulative. Suppose that public officials try to persuade people to engage in certain behavior with the help of relative risk information: “If you do not do X,

¹⁵¹ See Wilkinson, *supra* note.

¹⁵² Ruth Faden and Tom Beauchamp, *A History and Theory of Informed Consent* 354-68 (1986).

¹⁵³ Joseph Raz, *The Morality of Freedom* 377-79 (1986).

¹⁵⁴ See Wilkinson, *supra* note.

your chances of death from heart disease will triple!”¹⁵⁵ Suppose that for the relevant population, the chance of death from heart disease is very small – say, one in 50,000 – and that people are far more influenced by the idea of “tripling the risk” than they would be if they learned that they could increase a 1/50,000 risk to a 3/50,000 risk. On one view, the choice of the relative risk frame does not respect people’s decisionmaking capacities -- and it appeals directly to System 1.

115. Or suppose that choice architects are alert to the power of loss aversion,¹⁵⁶ and hence the use the “loss frame,” so as to trigger people’s concern about the risks associated with obesity and excessive energy consumption. They might deliberately choose to emphasize, in some kind of information campaign, how much people would lose from not using energy conservative techniques, rather than how much people would gain from using such techniques. On one view, the use of loss aversion is manipulative in the relevant sense: It does not (entirely, or quite) appeal to people’s deliberative processes but instead tries to trigger the negative feelings that are associated with losses.
116. But we have to be careful here, because a plausible understanding of manipulation might sweep up and perhaps condemn a great deal of conduct that is generally seen as unobjectionable, and reasonably so. It would be fussy, and to stringent, to condemn all such conduct, even if the word “manipulation” is reasonably applied to it.
117. Much of modern advertising is directed at System 1, with attractive people, bold colors, and distinctive aesthetics. (Consider advertisements for Viagra.) Cell phone companies, restaurants, and clothing stores use music and colors in a way that is designed to “frame” products in a distinctive manner. Doctors, friends, and family members (including spouses) often do something quite similar. Is romance an exercise is manipulation? Maybe so.¹⁵⁷ Is medical care? Is the use of social media? A great deal of conduct, however familiar, can be counted as manipulative in some relevant sense; but it would be extreme to condemn it for that reason.
118. On one view, nudges generally or frequently count as manipulative. Sarah Conly suggests that when nudges are at work, “Rather than regarding people as generally capable of making good choices, we outmaneuver them by appealing to their irrationality, just in more fruitful ways. We concede that people can’t generally make good

¹⁵⁵ Wilkinson, *supra* note, at 347, uses this example.

¹⁵⁶ See Eyal Zamir, *Law, Psychology, and Morality: The Role of Loss Aversion* (2014).

¹⁵⁷ I am aware of no detailed treatment of this question, but for relevant discussion, see Eric Cave, *Unsavory Seduction and Manipulation*, in *Manipulation*, *supra* note, at 176.

decisions when left to their own devices, and this runs against the basic premise of liberalism, which is that we are basically rational, prudent creatures who may thus, and should thus, direct themselves autonomously.”¹⁵⁸ (Conly is properly linking a concern with manipulation with the ideas of autonomy and dignity.)

119. This is a strong charge, and it is not fairly leveled against most kinds of nudges. Recall that many nudges are educative, and that many of them do not enlist or exploit System 1, or attempt in any way to “outmaneuver” people. But consider some testing cases, where the charge is not self-evidently misplaced.

(a) Choice architects might choose a graphic health warning, on the theory that an emotional, even visceral presentation might have significant effects.¹⁵⁹

(b) Recall that choice architects might be alert to *framing effects* and present information accordingly.¹⁶⁰

(c) They might be aware that a statement that a product is “90 percent fat-free” has a different impact from a statement that a product is “10 percent fat,” and they might choose the frame that has the desired effect.

(d) They might make a strategic decision about how to present social norms, knowing that the right presentation – for example, emphasizing behavior within the local community -- could have a large impact on people’s behavior.

(e) They might decide to list options – in a cafeteria or on a form – so as to make it more likely that people will make certain choices.

120. It is an understatement to say that none of these cases involves the most egregious forms of manipulation. There is no lying and no deceit. But is there an effort to subvert or to insult people’s decision-making powers? It is not absurd to say that at least some of these cases, the answer is yes.

121. I have said that government should be transparent about what it is doing. It should not hide its actions or its reasons for those actions. Does transparency rebut the charge of manipulation? Probably not. If government engages in egregious forms of manipulation, transparency is not a defense.¹⁶¹ A genuine insult to autonomy and dignity, in the

¹⁵⁸ Conly, *supra* note, at 30.

¹⁵⁹ See Christine Jolls, Product Warnings, Debiasing, and Free Speech: The Case of Tobacco Regulation, 169 *J Institutional and Theoretical Economics* 53 (2013).

¹⁶⁰ See Perspectives on Framing (Gideon Keren ed. 2010).

¹⁶¹ On the relationship between consent and manipulation, see Wilkinson, *supra* note, suggesting that consent can provide justification.

form of a subversion of people's decisionmaking capacities, does not become acceptable merely because people are allowed to know about it.¹⁶²

122. We could even imagine cases where full democratic control, alongside a high degree of transparency, are plainly insufficient to rebut that charge. Imagine that a democratic government adopted, freely and openly, a program of subliminal advertising – designed, let us say, for purposes of promoting public health (say, reducing smoking) rather than for illicit purposes. The program would be objectionable on the ground that it would be unacceptably manipulative, even if it was adopted openly and would produce good consequences.¹⁶³
123. Perhaps a graphic health warning could be counted as manipulative if it is designed to target people's emotions, rather than to inform them of facts.¹⁶⁴ But what if the warning is explained, in public, on exactly that ground? What if a warning is introduced and justified as effective, because it appeals to people's emotions, and thus saves lives? What if it is welcomed by the relevant population – say, smokers – for exactly that reason? Similar questions might be asked about strategic uses of framing effects, social norms, and order effects. T. M. Wilkinson convincingly argues that it is too crude to say that manipulation infringes upon autonomy, because “manipulation could be consented to. If it were consented to, in the right kind of way, then the manipulation would at least be consistent with autonomy and might count as enhancing it.”¹⁶⁵
124. We could understand consent as suggesting support from System 2, which might welcome a little manipulation (or possibly a lot) as a way of cabining the adverse effects of System 1 (recall present bias). To be sure, there are dangers in authorizing public officials to pursue this line of argument. But in certain contexts, the argument is more than plausible. Imagine, for example, a public education campaign that is designed to reduce the risks associated with texting while driving, or an effort to combat the use of drugs or to convince people to stay in school.

¹⁶² See note infra.

¹⁶³ I do not mean to resolve here the question whether a demonstration of good consequences could justify an admittedly manipulative action. In ordinary life, we could imagine a manipulative act – designed, say, to lead a child, a spouse, or a parent to take medicine – that would have an adequate justification.

¹⁶⁴ See note supra. There are also possible first amendment issues. Is it unconstitutional to require companies to include graphic warnings about the harms associated with their own products, if the requirement has a behavioral motivation, and is understood to be targeting System 1? See note supra.

¹⁶⁵ Wilkinson, supra note, at 345.

Many such campaigns are vivid and have an emotional component; they can be understood as efforts to combat self-control problems and to focus people on the long term.

125. If government is targeting System 1 – perhaps through framing, perhaps through emotionally evocative appeals – it may be responding to the fact that System 1 has already been targeted, and to people’s detriment. In the context of cigarettes, for example, it is plausible to say that a range of manipulations – including advertising and social norms – have influenced people to become smokers. If this is so, perhaps we can say that public officials are permitted to meet fire with fire. But some people might insist that two wrongs do not make a right – and that if the government seeks to lead people to quit, it must treat them as adults, and appeal to their deliberative capacities.
126. It is appropriate to conclude that even with full transparency, at least some degree of manipulation may be involved whenever a choice architect is targeting emotions or seeking a formulation that will be effective because of how it interacts with people’s intuitive or automatic thinking (System 1). But there are degrees of manipulation, and there is a large difference between a lie and an effort to frame an alternative in an appealing light. In ordinary life, we would not be likely to accuse our friends or loved ones of manipulation if they characterized one approach as favored by most members of our peer group, or if they emphasized the losses that might accompany an alternative that they abhor, or if they accompanied a description of one option with a frown and another with a smile.
127. Actions that are plausibly characterized as manipulative fall along a continuum, and if a doctor or a lawyer uses body language to support or undermine one or another alternative, it would be pretty fussy to raise objections about “subverting” or “perverting” the deliberative processes of a patient or client. It should be acknowledged that some nudges can be considered as manipulative within an ordinary understanding of that term. It should be emphasized that any action by government, including nudging, must meet a burden of justification. But when nudges fall within the periphery of the concept, when they have legitimate purposes, when they would be effective, and when they do not diverge from the kinds of influences that are common and unobjectionable in ordinary life, the burden of justification can often be met.

2. Manipulation, Self-Interest, and the Man Behind the Curtain

128. Thus far the discussion has been based on a particular understanding of manipulation – as subverting or perverting normal decisionmaking

capacities, or at least failing to appeal to them. But there are many other understandings. Let us explore some of them.

129. In an especially illuminating discussion, Anne Barnhill defines manipulation as “directly influencing someone’s beliefs, desires, or emotions, such that she falls short of ideals for belief, desire, or emotion in ways typically not in her self-interest or likely not in her self-interest in the present context.”¹⁶⁶ Because of its specificity and its focus on relevant ideals, the first part of the definition is especially useful. Note that “ideals” should not mean “ideal.” We do not think that belief, desire, or emotion need be based on some perfect (and dehumanized) process of formation. As Wilkinson notes, “the suggestion that manipulation is intentionally causing or encouraging people to decide in ways contrary to an ideally rational process cannot be correct. The root of the problem is that the ideally rational process is not in fact rational for us.”¹⁶⁷ Nonetheless, some processes of belief formation (for example) fall short not only of “the ideal” but of our own ideals. If so, we can complain of manipulation.
130. Barnhill’s formulation view fits well with Thomas Hill’s suggestion that “Manipulation, broadly conceived, can perhaps be understood as intentionally causing or encouraging people to make the decisions one wants them to make by actively promoting their making the decisions in ways that rational persons would not want to make their decisions.”¹⁶⁸ Much of the foregoing analysis could proceed under Barnhill’s or Hill’s formulation; it would not be fundamentally changed. Of course it would be necessary to be clear on what it means to “fall short of” the relevant ideals.
131. The second feature of Barnhill’s definition builds into the definition an inquiry into whether the influence is in the person’s self-interest. This aspect of the definition should probably be discarded. Suppose that a government, or a private company, influences people’s beliefs in a way that falls far short of ideals for belief formation. For example, people might be presented with a distorted or partial picture of the situation (involving, say, the risks associated with obesity). Even if the influence is in people’s self-interest – because it leads them to alter their conduct in a way that lengthens their lives – we might insist that they have been manipulated. A subliminal advertisement counts as manipulative even if people are far better off as a result of being exposed to it.

¹⁶⁶ Anne Barnhill, What is Manipulation? in *Manipulation: Theory and Practice* 50, 72 (Christian Coons and Michael Weber eds. 2014).

¹⁶⁷ Wilkinson, *supra* note, at 350.

¹⁶⁸ Thomas Hill, *Autonomy and Self-Respect* 33 (1991).

132. In ordinary usage, manipulation is sometimes said to exist when *someone tries to alter behavior in a covert way, by hiding, or at least not revealing, an important aspect of the relevant interaction.*¹⁶⁹ Recall that a lie is a defining example of manipulation. If you tell people that they should eat a certain kind of cereal because it is healthy to do so, and if the facts are otherwise, then you are manipulating them. Deceptive behavior counts as manipulative as well, even if no one has actually spoken falsely. If you imply that certain food is unhealthy to eat, when it is not, you are manipulating people's behavior. In accordance with the previous discussion, it is useful to ask whether choice architects are bypassing or subverting ordinary decisionmaking capacities, but the idea of hiding, or not revealing, an important aspect of the interaction captures something significant about the idea of manipulation.
133. On this account, an action counts as manipulative if it *lacks transparency* – if the role or the motivation of the choice architect is hidden or concealed. In the pivotal scene in *The Wizard of Oz*, the Wizard says, “Pay no attention to the man behind the curtain.” The man behind the curtain is of course a mere human being who is masquerading as the great Wizard – and who is both claiming far more authority than he deserves and who is designing social situations in a way that hides features that, if revealed, could alter people's judgments and choices. If choice architects conceal their own role, it seems fair to charge them with being manipulative.
134. We have seen most nudges are not manipulative in any relevant sense. But to the extent that some of them are, we can imagine a reasonable objection or concern, whose force depends on the degree of the manipulation. Perhaps we could have an absolute or near-absolute taboo on lying or deception on government's part, for welfarist or nonwelfarist reasons.¹⁷⁰ But for reasons discussed above, we should be more lenient toward emotional appeals and framing. Because government always faces a burden of justification, one question is whether such approaches produce significant welfare gains. If a graphic health warning saves many lives, it is unacceptable if and because it can be counted as a (mild) form of manipulation? A welfarist would want to make an all-things-considered judgment about the welfare consequences.
135. It is true that some people, focused on autonomy as an independent good, would erect a strong and perhaps conclusive presumption against

¹⁶⁹ Barnhill, *supra* note.

¹⁷⁰ Under true emergency conditions – for example, when national security is genuinely threatened – it is possible that this constraint can be overcome.

defining, clear, or egregious cases of manipulation.¹⁷¹ But at least in general, the modest forms discussed here strain the boundaries of the concept, and it would be odd to rule them off-limits.

F. Learning

136. Choice-making is a muscle, and the ability to choose can be strengthened through exercise. If nudges would make the muscle atrophy, we would have an argument against them. We could imagine an ethical objection that would contend that some nudges do not allow people to build up their own capacities, and might even undermine their incentive to do so.¹⁷² Here too, it is necessary to investigate the particulars – the kinds of nudges and choice architecture that are involved.
137. Active choosing and prompted choice hardly impede learning. Nor do information and reminders. On the contrary, they promote learning. Here the evidence is compelling: Nudges of this kind exercise the choice-making muscle, rather than the opposite.¹⁷³
138. With respect to learning, a potential problem comes from default rules. It is possible to say that active choosing is far better than defaults, simply because choosing may promote learning. Consider, for example, the question whether employers should ask employees to make active choices about their retirement plans, or whether they should instead default people into plans that fit their situations. The potential for learning might well count in favor of active choosing.¹⁷⁴ If people are defaulted into certain outcomes, they do not add to their stock of knowledge, and that may be a significant lost opportunity.
139. The argument for learning depends on the setting. (Recall the earlier discussion of educative nudges.) For most people, it is not important to become experts in the numerous decisions that lead to default settings in cell phones, and hence the use of such settings is not objectionable. The same point holds in many other contexts in which institutions rely on defaults rather than active choosing. To know whether choice architects should opt for active choosing, it is necessary to explore whether the context is one in which it is valuable, all things considered, for choosers to acquire a stock of knowledge.

¹⁷¹ Cf. White, *supra* note.

¹⁷² See Gigerenzer, Risk Savvy, *supra* note.

¹⁷³ See Benjamin York and Susanna Loeb, One Step at a Time: The Effects of an Early Literacy Text Messaging Program for Parents of Preschoolers (2014), available at <http://www.nber.org/papers/w20659>

¹⁷⁴ See Cass R. Sunstein, Choosing Not To Choose (forthcoming 2015).

G. Biased Officials

140. Choice architects are emphatically human, and fully subject to behavioral biases; they are often unreliable. The growing field of *behavioral public choice* draws on this point to offer an account of official error.¹⁷⁵ It is reasonable to object to some nudges, and to some efforts to intervene in existing choice architecture, on the ground that the choice architects might blunder.¹⁷⁶ They might lack important information (the knowledge problem). They might be biased, perhaps because their own parochial interests are at stake (the public choice problem). They might themselves display behavioral biases – suffering, for example, from present bias, optimistic bias, or probability neglect. In a democratic society, public officials are responsive to public opinion, and if the public is mistaken, officials might be mistaken as well.
141. It is unclear whether and to what extent this objection is a distinctly ethical one, but it does identify an important cautionary note. One reason for nudges, as opposed to mandates and bans, is that choice architects may err.¹⁷⁷ No one should deny that proposition, which argues in favor of choice-preserving approaches. If choice architects blunder, at least it can be said that people are entitled to go their own way. And if we emphasize the risk of official error, we might want to avoid public officials to avoid nudges and choice architecture as well.
142. The initial response to this objection should be familiar: Choice architecture is inevitable. When choice architects act, they alter the architecture; they do not create an architecture where it did not exist before. A certain degree of nudging, from the public sector, cannot be avoided, and there is no use in wishing it away. Nonetheless, choice architects who work for government might decide that it is best to rely on free markets, and to trust in invisible hand mechanisms. If so, they would select (or accept) choice architecture that reflects those mechanisms.
143. This idea raises many conceptual and empirical puzzles, to which I have referred above, and which I will not engage in detail here. The question is whether it is so abstract, and so rooted in dogmas, that it ought not to

¹⁷⁵ For one example, see Timur Kuran and Cass R. Sunstein, *Availability Cascades and Risk Regulation*, 51 *Stan. L. Rev.* 683 (1999); for an overview, see Jan Schnellenbach and Christian Schubert *Behavioral Public Choice: A Survey* (2014), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2390290

¹⁷⁶ Rebonato, *supra* note.

¹⁷⁷ See Cass R. Sunstein, *Nudges vs. Shoves*, 127 *Harv. L. Rev. Forum* 210 (2014).

command support. To be sure, free markets have many virtues. But in some cases, disclosure, warnings, and reminders can do far more good than harm.¹⁷⁸ As we have seen, active choosing is sometimes inferior to default rules. Someone has to decide in favor of one or another, and in some cases, that someone is inevitably the government. It is true that distrust of public officials will argue against nudging, at least where it is avoidable, but if it is dogmatic and generalized, such distrust will likely produce serious losses in terms of both welfare and freedom.

VI. What Do People Think?

144. What do people actually think about nudging and choice architecture? I conclude with three empirical findings. These findings cannot, of course, dispose of the ethical questions. The question is how to answer those questions in principle, and findings about people's answers are not decisive. But they do help to illuminate matters, and for two different reasons. The first is that in a democratic society, it is important to attend to what people actually think. If people have strong objections, democratic governments should hesitate before proceedings, and in any case they probably will do so (because of electoral self-interest). The second reason is that people's judgments provide relevant information about to think about the ethical issues even if that information is not conclusive. As we shall see, people agree that the nature of the particular nudge in question is relevant to the ethics of nudging.

A. The Effects of Transparency About Nudging

145. Here is a preliminary question, relevant to ethical questions: If people are explicitly informed that they are being nudged, does their behavior change? An important study by George Loewenstein, Cindy Bryce, and David Haggman offers the following finding, at least in one context: *When people are specifically informed that a default rule has been put in place, and that it might be otherwise, that information has essentially no effect on what people do.*¹⁷⁹ At least in the context of end-of-life care, the effect of a default is not weakened when people are told that a default was chosen because it is usually effective.
146. Here is the disclosure: "The specific focus of this research is on 'defaults' – decisions that go into effect if people don't take actions to

¹⁷⁸ See Oren Bar-Gill, *Seduction by Contract* (2012).

¹⁷⁹ See George Loewenstein, Cindy Bryce, David Haggmann & Sachin Rajpal, *Warning: You Are About To Be Nudged* 17 (Mar. 29, 2014) (unpublished working paper) (<http://ssrn.com/abstract=2417383>) ("[I]nforming people about default interventions in advance . . . does not significantly diminish the impact of defaults on expressed preferences in advance directives.").

do something different. Participants in this research project have been divided into two experimental groups.” Having received this information, participants were told, “If you have been assigned to one group, the Advance Directive you complete will have answers to questions checked that will direct health care providers to help relieve pain and suffering even it means not living as long. If you want to choose different options, you will be asked to check off different option and place your initials beside the different option you select.” Participants were also told that “if you have been assigned to the other group, the Advance Directive you complete will have answers to questions checked that will direct health care providers to prolong your life as much as possible, even if it means you may experience greater pain and suffering.” Notably, this information did not affect people’s ultimate choices.

147. It would be interesting to know if the results would have been different if people had been told something like this: “Default rules often have significant effects on behavior, because of the force of inertia, and because people often think that such rules reflect what most people do.” It would also be interested to know if the setting of end-of-life care is distinctive in this regard, and if larger effects, from the design in the Loewenstein experiment, would be found in other contexts. But it is not unreasonable to speculate that at least in many contexts, disclosure that a default rule is in place, and could be otherwise, would not much affect outcomes.
148. A warning, a prompted choice, or a reminder is not exactly mysterious; people are certainly aware that the goal is to nudge. We can imagine situations in which warnings and reminders backfire; perhaps people will do exactly what they are warned or reminded not to do. The idea of “reactance” points to this possibility,¹⁸⁰ and it must be taken into account. But in the absence of unusual circumstances, warnings and reminders are likely to have their intended effect.

B. Politics Matters

149. It is reasonable to speculate that political values greatly matter to people’s assessment of nudges.¹⁸¹ In other words, people often have no

¹⁸⁰ See Sharon Brehm and Jack Brehm, *Psychological Reactance: A Theory of Freedom and Control* (1981); Louisa Pavey & Paul Sparks, *Reactance, Autonomy and Paths to Persuasion: Examining Perceptions of Threats to Freedom and Informational Value*, 33 *MOTIVATION & EMOTION* 277 (2009).

¹⁸¹ David Tannenbaum et al., *On the Misplaced Politics of Behavioral Policy Interventions* (unpublished manuscript 2014).

particular view on nudges as such. Their evaluations turn on whether they like the politics of the particular nudge.

150. Casual observation suggests that many people like or dislike nudges because of their perceived merits, not because they are nudges. More systemic evidence supports this view.¹⁸² The result is “partisan nudge bias,” as “people find nudges more ethically problematic when they are applied to policy objectives they oppose, or when applied by policymakers they oppose, while they find the same nudges more acceptable when they are applied to political objectives they support or by policymakers they support.”¹⁸³ People who are right-of-center are less likely to approve of nudges that seem to have a distinctive left-of-interest motivation; they are more likely to approve of nudges if they are right-of-center. People who are left-of-center show the same pattern.
151. There is a clear implication here for the political economy of nudging: People’s judgments are likely to be, in significant part, an artifact of their substantive judgments about the direction in which people are being nudged. True, we should not take this point too far. Those who support an incumbent president would be likely to object if he imposed a nudge that entrenched himself (as, for example, through a system of default voting). In egregious cases of manipulation, citizens of a free society (or even an unfree one) might well be outraged whatever they think of the underlying substance. But within certain limits, political assessments are likely to reflect judgments about that substance.

C. Nudging System 1

152. Consider finally an intriguing finding: People care whether nudges are overt or covert – but not always, and not always a great deal.¹⁸⁴ The terms “overt” and “covert” are not self-defining, but we can imagine a difference between nudges that appeal to deliberative capacities (System 2) or to unconscious processing or intuitions (System 1). Notice, for example, the difference between two scenarios.¹⁸⁵

(a) The new design works like this—with every annual salary increase you are provided information in the form of a series of icons representing tropical beaches that shows how much extra leisure you are likely to be able to afford during your retirement by

¹⁸² Id.

¹⁸³ Id.

¹⁸⁴ For relevant discussion, see Gidon Felsen et al., *Decisional Enhancement and Autonomy: Public Attitudes Toward Overt and Covert Nudges*, 8 *Judgment and Decision Making* 203 (2012).

¹⁸⁵ Id.

investing different percentages of your increased salary; larger investments now translate into more retirement savings later. You can still choose to keep the entire salary increase instead of investing it, but the information provided results in a subconsciously-driven bias towards investment; in other words, the decision to invest is made more likely as a result of subconscious deliberation. Studies have shown that implementing this policy leads to an increase in retirement savings.

(b) The new design works like this—with every annual salary increase you are provided information in the form of a detailed table of your earnings that shows how much extra money you are likely to have during your retirement by investing different percentages of your increased salary; larger investments now translate into more retirement savings later. You can still choose to keep the entire salary increase instead of investing it, but the information provided results in a consciously-driven bias towards long-term investment; in other words, the decision to invest is made more likely as a result of conscious deliberation. Studies have shown that implementing this policy leads to an increase in retirement savings.

153. Such questions allow for a test of this hypothesis, connected with the earlier discussion of manipulation: *people object to nudges that are covert in the sense that they appeal to unconscious or subconscious processes.* There is evidence that people do indeed show a preference for nudges that do not have that characteristic.¹⁸⁶ One reason is that such appear to be less manipulative, because they engage higher-order thinking. Another reason may be that when nudges appeal to unconscious or subconscious processes, they might seem to offend individual dignity.
154. Notably, however, the difference in reactions is only moderate, and when people believe that some kind of behavioral bias – such as a self-control problem -- is genuinely responsible for welfare losses, they become more receptive to nudges that target unconscious or subconscious processes.¹⁸⁷ As we have seen, people’s judgments about the ethical questions cannot resolve the ethical questions. But it is illuminating to see both that people show a preference for nudges that cannot be categorized as manipulative, and that they show a degree of receptivity to arguably manipulative ones.

VII. Conclusion

¹⁸⁶ Id.

¹⁸⁷ Id.

155. It is pointless to object to nudges and choice architecture as such. Human beings cannot live in a world without them. Spontaneous orders have many virtues, but they themselves nudge. Whether or not they are associated with liberty, properly conceived, they create multiple forms of choice architecture. Even the most minimal government must nudge, and must create choice architecture of many different kinds. A modest regulatory state will influence people's decisions even if it seeks not to do so. Consider the effects of default rules, of the sort that are pervasive in the law of property, contract, and tort.
156. The modern regulatory state imposes numerous mandates and bans, and some of them are properly characterized as paternalistic. Consider the requirement that people obtain prescriptions before using certain medicines, or fuel economy and energy efficiency rules, or occupational safety and health law; all these, and many others, have paternalistic features. Paternalistic mandates and bans are subject to obvious ethical concerns, many of them identical to those explored here.¹⁸⁸ Because nudges preserve freedom of choice, those concerns are weakened.
157. Nonetheless, any changes in choice architecture, including those that preserve freedom, can run into serious and even convincing ethical objections – most obviously, where the underlying goals are illicit. But where the goals are legitimate, nudges are less likely to run afoul of ethical constraints, not least because and when they promote informed choices (as in the case of reminders). Transparency and public scrutiny are important safeguards, especially when public officials are responsible for nudges and choice architecture. Nothing should be hidden or covert.
158. Nonetheless, some imaginable nudges are objectionable, even when legitimate goals are involved, even when freedom of choice is preserved, and even in the face of full transparency. Most important, some nudges can be counted as forms of manipulation, raising objections from the standpoint of both autonomy and dignity.
159. That is a strong point against them. Even when nudges target System 1, it might well strain the concept of manipulation to categorize them as such (consider a graphic warning). The concept of manipulation has a core and a periphery; some nudges fit within the core, others within the periphery, and others outside of both.

¹⁸⁸ A valuable discussion is Conly, *supra* note.

160. Many nudges, and many changes in choice architecture, are not merely permissible on ethical grounds; they are actually required. On grounds of welfare, the point should be straightforward; much nudging promises to increase social welfare. But the point holds for autonomy, dignity, and self-government as well.
161. The history of freedom-respecting nations is full of changes in choice architecture that have permitted them to move further in the direction of their highest ideals. It should go without saying that those ideals have yet to be fully realized. In moving closer to them, new nudges, and new forms of choice architecture, will prove indispensable.