

NUDGING VIRTUE

KIRAN IYER*

I. THE ETHICS OF NUDGING

Governments around the world are using nudges, “interventions that steer people in particular directions but that also allow them to go their own way,”¹ to influence behavior. Examples include automatic enrollment into retirement savings plans, graphic warnings about the risks of smoking, and informing people about how their energy consumption compares to their neighbors. Drawing on behavioral science research, nudgers seek to help people make better choices by altering the “choice architecture” within which they make decisions.² Nudges differ from other policies that seek to affect individual choices, such as mandates and subsidies, in that they do not impose significant material incentives or disincentives on choosers.³ Since the United Kingdom launched its Behavioral Insights Team in 2010, governments in the United States, Germany, and Australia have moved to set up their own “nudge units” aimed at applying behavioral insights to policymaking.⁴ Underscoring nudging’s policy significance, President Obama directed federal agencies in September 2015 to “identify policies, programs, and operations where applying behavioral science insights may yield substantial improvements in public welfare, program outcomes, and program cost effectiveness.”⁵

Nudging has been heralded as offering a “Third Way” for lawmakers seeking to promote desirable behaviors such as saving for retirement, charitable giving, and energy conservation without compromising freedom of choice.⁶ However, nudging remains controversial. Existing objections

* Assistant Public Defender - Appellate Division, Maryland Office of the Public Defender. This paper was prepared with the support of the Harvard Law School Summer Academic Fellowship Program. I am grateful to Cass Sunstein for supervising this project, and to Jeffrey Goldsworthy, Justin Oakley, Lawrence Solum, Damien Bruckard, James Wilson, and Heather Anderson for their comments on earlier drafts. All errors remain my own.

1 Cass R. Sunstein, *The Ethics of Nudging*, 32 YALE J. ON REG. 413, 417 (2015). Those nudges that aim to steer people’s choices in ways that improve their own welfare have been justified by Sunstein and Thaler as a form of “libertarian paternalism”: Cass R. Sunstein & Richard Thaler, *Libertarian Paternalism is not an Oxymoron*, 70 U. CHI. L. REV. 1159, 1162 (2003).

2 RICHARD H. THALER & CASS R. SUNSTEIN, NUDGE: IMPROVING DECISIONS ABOUT HEALTH, WEALTH, & HAPPINESS 6 (2008).

3 Sunstein, *supra* note 1, at 417.

4 For evidence of the growing influence of behavioral science findings on policy around the world, see PETE LUNN, OECD, REGULATORY POLICY & BEHAVIORAL ECONOMICS 17-25 (2014).

5 Exec. Order No. 13,707, 80 Fed. Reg. 56,365 (Sept. 15, 2015).

6 Thaler & Sunstein, *supra* note 2, at 252-253. For a description of behaviorally informed interventions used successfully by the United Kingdom’s Behavioral Insights Team, see DAVID HALPERN, INSIDE THE NUDGE UNIT (2015). However, nudges are infrequently subjected to a full social welfare evaluation, meaning that their welfare benefits may sometimes be overstated: see Hunt Allcott

correspond to the two most prominent traditions in Western normative ethics: consequentialism and deontology. Consequentialists argue that nudging fails to improve individual or social welfare because: (1) choice architects cannot determine what people really want;⁷ (2) nudging fails to account for the value individuals attribute to unsteered choices;⁸ (3) nudges impede learning;⁹ (4) choice architects tend to err, due to behavioral biases, or misuse behavioral findings;¹⁰ and (5) nudges backfire by prompting reactance.¹¹ Deontologists contend that nudges impermissibly (1) intrude on autonomy,¹² (2) compromise dignity,¹³ and/or (3) manipulate behavior.¹⁴

This paper offers a fresh analysis of the ethics of nudging, drawing from the third major Western ethical tradition—virtue ethics. Virtue ethics, in the form outlined in this paper, is distinctive from the other approaches because it makes “essential reference” to character and virtue in the justification of right action, rather than treating consequences (consequentialism) or conformity to moral duties or rules (deontology) as dispositive.¹⁵ It shifts the focus of ethical analysis by assessing whether nudges are consistent with a thick account of human flourishing. When developed into a framework for evaluating policy, it enables us to consider some crucial questions about nudging: What impact do nudges have on the development of character? Under what circumstances do nudges promote human flourishing? How could virtue-based concepts inform the decision to nudge? The present debate about nudging has been impoverished by the failure to confront these questions; this paper fills that gap.

The paper proceeds as follows: Part II identifies the specific nudges considered in this paper. Part III outlines a neo-Aristotelian account of virtue ethics and translates it into a framework for policy evaluation. It also responds to some of the obvious objections to evaluating nudging through

& Judd Kessler, *The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons 2–5* (Nat'l Bureau of Econ. Research, Working Paper No. 21671, 2015), <http://www.nber.org/papers/w21671>.

⁷ This claim encompasses a series of objections, including that choice architects do not have enough information about people's ends and means, and that people's preferences adapt in a manner that renders the search for preferences unfeasible: see Riccardo Rebonato, *A Critical Assessment of Libertarian Paternalism*, J. CONSUMER POL'Y 357, 363–365, 376–379 (2014); CASS R. SUNSTEIN, WHY NUDGE? 91–92 (2014).

⁸ Rebonato, *supra* note 7, at 382.

⁹ Anne van Aaken, *Judge the Nudge: In Search of the Legal Limits of Paternalistic Nudging in the EU*, in NUDGE & THE LAW: A EUROPEAN PERSPECTIVE 83, 95 (Alberto Alemanno & Anne-Lise Sibony eds., 2015).

¹⁰ W. Kip Viscusi & Ted Gayer, *Behavioral Public Choice: The Behavioral Paradox of Government Policy*, 38 HARV. J.L. & PUB. POL'Y. 973, 977–980 (2015).

¹¹ For some suggestive evidence of reactance, see Ayala Arad & Ariel Rubinstein, *The People's Perspective on Libertarian-Paternalistic Policies* (July 2015), <http://www.tau.ac.il/~aradayal/LP.pdf>.

¹² Rebonato, *supra* note 7, at 381–383.

¹³ Jeremy Waldron, *It's All for Your Own Good*, N.Y. REV OF BOOKS, Oct. 9, 2014, <http://www.nybooks.com/articles/2014/10/09/cass-sunstein-its-all-your-own-good/>.

¹⁴ For different versions of this critique, see MARK D. WHITE, THE MANIPULATION OF CHOICE: ETHICS & LIBERTARIAN PATERNALISM 91 (2013); SARAH CONLY, AGAINST AUTONOMY: JUSTIFYING COERCIVE PATERNALISM 30 (2012). See also Cass R. Sunstein, *Fifty Shades of Manipulation*, 1 J. MKT. BEHAV. 214 (2016).

¹⁵ Roger Crisp & Michael Slote, *Introduction*, in VIRTUE ETHICS 2–3 (Roger Crisp & Michael Slote eds., 1997).

the prism of virtue ethics. Part IV provides a virtue-based analysis of the selected nudges. Part V concludes by outlining some guiding principles for lawmakers.

II. SCOPE OF THIS PAPER

Nudging encompasses a broad range of policies directed at influencing choices without imposing significant material costs on choosers:

The most obvious nudges consist of default rules, which establish what happens if people do nothing at all. 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 strategies (through which people agree, in advance, to a particular course of conduct, such as a smoking cessation program).¹⁶

As should be evident from the diversity of this list, nudging is not *per se* ethical or unethical; what matters is the specific nudge being considered and the context in which that nudge is deployed.¹⁷ This paper will scrutinize six prominent nudging techniques. The first three have been used to promote sustainable energy consumption; the last three to combat smoking:

- A. *Default rules*: automatic enrolment in green energy programs
- B. *Active choosing*: required choosing between different energy programs
- C. *The use of social norms*: informing consumers of their energy consumption relative to their neighbors
- D. *Factual disclosure*: factual information about the risks of smoking
- E. *Graphic warnings*: unpleasant images on cigarette packets
- F. *Precommitment strategies*: smoking cessation programs

This selection enables a comparison between nudges that differ across the three domains identified by Sunstein:¹⁸

- A. *System 1 nudges* (nudges that seek to activate or rely upon our automatic, intuitive judgments, such as default rules, graphic warnings, and the use of social norms) and *System 2 nudges*

¹⁶ Sunstein, *supra* note 1, at 424.

¹⁷ Even the most ardent libertarian would be unlikely to oppose the simplification of (necessary) government forms.

¹⁸ Sunstein, *supra* note 1, at 426–428.

(nudges that seek to enhance the role of our deliberative capacities, such as active choosing, factual disclosure, and precommitment);¹⁹

- B. *Paternalistic nudges* (nudges that seek to influence people's choices in their own interests, such as those aimed at discouraging smoking) and market failure nudges (such as those deployed to reduce externalities caused by dirty energy consumption); and
- C. *Educative nudges* aimed at increasing knowledge or understanding (like factual disclosure and active choosing) and *non-educative nudges* (like graphic warnings and default rules).

III. VIRTUE-BASED FRAMEWORK FOR POLICY EVALUATION

Virtue ethics is “both an old and a new approach to ethics,”²⁰ harkening back in Western philosophy to Socrates, Plato, and Aristotle, but finding contemporary significance in the work of modern moral philosophers.²¹ Since its revival in the second half of the twentieth century,²² it has challenged the primacy of the dominant traditions in Western normative ethics—consequentialism and deontology. As a first approximation, virtue ethics differs from those theories because it “emphasizes the virtues, or moral character.”²³ When asked whether an action is right, virtue ethicists consider “what a virtuous agent would, characteristically, do in the circumstances,” rather than whether that action would have good consequences (consequentialism) or conform to moral duties or rules (deontology).²⁴ The “common ground” within virtue ethics has been more fully sketched by Nussbaum:

- A. Moral philosophy should be concerned with the agent, as well as with choice and action.
- B. Moral philosophy should therefore concern itself with motive and intention, emotion and desire: in general, with the character of the inner moral life, and with settled patterns of motive, emotion, and reasoning that lead us to call someone a person of a certain sort (courageous, generous, moderate, just, etc.).
- C. Moral philosophy should focus not only on isolated acts of choice, but also, and more importantly, on the whole course of the agent's moral life, its patterns of commitment, conduct, and also passion.²⁵

¹⁹ The distinction between System 1 and System 2 families of cognitive operations was popularized by DANIEL KAHNEMAN, *THINKING FAST AND SLOW* 20–21 (2011).

²⁰ ROSALIND HURSTHOUSE, *ON VIRTUE ETHICS* 1 (1999).

²¹ Some of the leading early texts included ALASDAIR MACINTYRE, *AFTER VIRTUE* (1981); PHILIPPA FOOT, *VIRTUES AND VICES AND OTHER ESSAYS IN MORAL PHILOSOPHY* (1978); and BERNARD WILLIAMS, *ETHICS AND THE LIMITS OF PHILOSOPHY* (1985)..

²² The revival of virtue ethics is usually traced to a 1958 article by Elizabeth Anscombe attacking the other ethical theories and calling for a return to ancient notions of virtue: G.E.M. Anscombe, *Modern Moral Philosophy*, 33 *PHIL.* 1 (1958).

²³ Hursthouse, *supra* note 20, at 1.

²⁴ *Id.* at 17.

²⁵ Martha C. Nussbaum, *Virtue Ethics: A Misleading Category?*, 3 *J. OF. ETH.* 163, 170 (1999).

Modern consequentialists and deontologists have acknowledged the moral significance of the agent's character.²⁶ However, as Crisp and Slote explain, virtue ethics remains a conceptually distinct theory because it:

... makes essential reference to the rationality of virtue itself. Thus, for example, the real reason why I should not lie to you is not that it is against the moral law, nor that it is likely not to maximize well-being, but because it is *dishonest*. The notions of virtue, then, are more basic than the notions at the heart of utilitarian and Kantian theory ... it is characteristic of modern virtue ethics that it puts primary emphasis on aretaic or virtue-centered concepts rather than deontic or obligation-centered concepts.²⁷

Virtue ethics encompasses a range of different perspectives. Theorists who subscribe to some form of virtue ethics differ with regard to the following questions: (1) What are the virtues?²⁸ (2) Are the virtues universal or culturally relative?²⁹ (3) Can the virtues be harmonized?³⁰ (4) Are the virtues intrinsically good, or are they valuable because they are necessary to the good life (*eudaimonism*) or realize our essential properties as human beings (*perfectionism*)?³¹ (5) To what extent can we exercise reason to guide our behavior and determine what is virtuous?³² (6) How do we become virtuous?³³ (7) Should virtue ethics be translated into virtue politics, and, if so, how should that occur?³⁴

This paper does not defend virtue ethics against the other ethical theories, and brackets many of its internal conflicts. Instead, it sets out a model theory, adopting a neo-Aristotelian account of virtue ethics inspired

26 See, e.g., BARBARA HERMAN, *THE PRACTICE OF MORAL JUDGEMENT* (1993) (emphasizing the common ground between virtue ethics and Kantianism); Philippa Foot, *Utilitarianism and the Virtues*, 94 *MIND* 196 (1985) (arguing that utilitarianism and virtue ethics can be reconciled). Claims of this type have been reinforced by the discussions of virtue in the works of Immanuel Kant, John Stuart Mill, and Henry Sidgwick: see Nussbaum, *supra* note 25, at 165–167.

27 Crisp & Slote, *supra* note 15, at 3.

28 See MACINTYRE, *supra* note 21, at 181–203 (comparing the accounts of virtue in the works of Homer, Aristotle, and Aquinas). For a selection of articles outlining the conceptions of virtue advanced by Plato, Aristotle, Confucius, Hume, and Nietzsche, see *THE ROUTLEDGE COMPANION TO VIRTUE ETHICS* (Lorraine Besser-Jones & Michael Slote eds., 2015).

29 See generally Martha Nussbaum, *Non-Relative Virtues: An Aristotelian Approach*, in *ETHICS: ESSENTIAL READINGS IN MORAL THEORY* 446 (George Sher ed., 2012) (opposing the turn towards relativism within contemporary virtue ethics); compare MACINTYRE, *supra* note 21.

30 See generally Daniel Russell, *Phronesis and the Virtues*, in *THE CAMBRIDGE COMPANION TO ARISTOTLE'S NICOMACHEAN ETHICS* 203, 203 (Ronald Polansky ed., 2014) (comparing the “reciprocity of the virtues” thesis (having any virtue is to have them all) with the “unity of the virtues” thesis (all the virtues are the same)).

31 See Crisp & Slote, *supra* note 15, at 21–23 (comparing Hursthouse's *eudaimonist* approach with Slote's emphasis on the intrinsic moral character of certain virtues). See also THOMAS HURKA, *VIRTUE, VICE, AND VALUE* (2001) (defending a perfectionist account of virtue).

32 See Nussbaum, *supra* note 23, at 168–169 (comparing anti-Utilitarian approaches to virtue ethics, that seek to expand the role of reason and deliberation, with anti-Kantian approaches, which give greater prominence to sentiments and habits).

33 For a range of perspectives on cultivating virtue, see *CULTIVATING VIRTUE: PERSPECTIVES FROM PHILOSOPHY, THEOLOGY, & PSYCHOLOGY* (Nancy Snow ed., 2015).

34 See generally Crisp & Slote, *supra* note 15, at 24–25; Catherine H. Zuckert, *Aristotelian Virtue Ethics and Modern Liberal Democracy*, 68 *REV. OF METAPHYS.* 61 (2014). For one approach to translating Aristotle's ideas into political philosophy, see MARTHA C. NUSSBAUM, *FRONTIERS OF JUSTICE: DISABILITY, NATIONALITY, SPECIES MEMBERSHIP* (2006).

by Hursthouse³⁵ and Annas³⁶ and drawing guiding principles for lawmakers from Solum's "aretaic theory of legislation."³⁷ That model theory is captured in the following propositions:

- A. *The ultimate end of human life is eudaimonia. Eudaimonia is the end that we choose for its own sake, and for the sake of which we choose all other ends.*³⁸ *Eudaimonia* has been variously translated as happiness, well-being, and flourishing.³⁹ However, *eudaimonia* is "avowedly, a moralized, or 'value-laden' concept of happiness, something like 'true' or 'real' happiness or 'the sort of happiness worth seeking or having.'"⁴⁰ Happiness, in this sense, has subjective and objective components:

Happiness has the role of being, for each person, *your* happiness, the way *you* achieve living your life well. It is not some plan imposed on you from outside ... At the same time it is not just anything you want it to be. There are better and worse ways of seeking happiness, for there are clearly better and worse ways of organizing your goals and aims in life, and of seeking to live a life that achieves them overall.⁴¹ (Emphasis added).

- B. *Human beings achieve eudaimonia by developing and exercising the virtues.* As Aristotle explained, the "best life for a human being ... consists in the exercise of the virtues."⁴² This is based on the naturalistic view that "the characteristic activity of human beings is the exercise of our rational capacity, and only by living virtuously is our rational capacity to guide our lives expressed in an excellent way."⁴³ True happiness, on this account, is not constituted by the "stuff you have, or whether you are beautiful, healthy, powerful, or rich."⁴⁴ Rather, it depends on how you "deal with the material of your life."⁴⁵ Although favorable material conditions are not constitutive of the human good, they do play an instrumental role in affecting the human capacity to exercise the virtues.⁴⁶ For

35 Hursthouse, *supra* note 20. Appiah describes Hursthouse's account of virtue ethics as "representative, because it succeeds in capturing elements that are shared by many of the doctrinal variants in circulation": KWAME ANTHONY APPIAH, EXPERIMENTS IN ETHICS 36 (2008).

36 JULIA ANNAS, INTELLIGENT VIRTUE (2011).

37 Lawrence B. Solum, Virtue as the End of Law: An Aretaic Theory of Legislation, JURIS. (forthcoming 2016).

38 ARISTOTLE: NICOMACHEAN ETHICS 10 (Roger Crisp ed., 2014).

39 Each of these translations has its limitations: see Richard Kraut, *Two Conceptions of Happiness*, 88 PHIL. REV. 167, 167-169 (1979). However, for ease of reading, these terms will be used interchangeably in what follows.

40 Rosalind Hursthouse, *Virtue Ethics*, STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Jul. 18, 2003), <http://plato.stanford.edu/entries/ethics-virtue/>.

41 Annas, *supra* note 36, at 126.

42 Crisp & Slote, *supra* note 15, at 2.

43 Justin Oakley, *A Virtue Ethics Approach*, in A COMPANION TO BIOETHICS 91, 96 (Helga Kuhse & Peter Singer eds., 2nd ed. 2009).

44 Annas, *supra* note 36, at 129.

45 *Id.*

46 See generally Martha Nussbaum & David Charles, *Nature, Function, and Capability: Aristotle on Political Distribution*, 6 OX. STUD. PHIL. 145 (1988) (explaining Aristotle's view that the right amount of an external good is that which makes human functioning best).

example, a person who faced serious deprivation as a child may find it more difficult to exercise self-control over her desires and live a flourishing life.⁴⁷

- C. *Virtue involves “appropriate functioning” in each sphere of human experience.*⁴⁸ The virtues are the specification of whatever it means to choose well within a “sphere of human experience that figures in more or less any human life.”⁴⁹ Those spheres include bodily appetites and their pleasures, the distribution of limited resources, social relations, giving and retaining money, and the planning of one’s life and conduct.⁵⁰ In all of these spheres, across most (if not all) cultures, humans must make choices which may be considered better or worse. The virtues are those traits that reflect success within each sphere.⁵¹ They include moral virtues (honesty, justice, generosity, courage, loyalty, temperance, benevolence, and kindness)⁵² and intellectual virtues (practical wisdom and theoretical wisdom).

Acting from virtue is demanding – it involves reasoning, feeling, and acting in the ethically right way, rather than merely having a tendency to behave in that way.⁵³ As Hursthouse explains, a virtue is a disposition:

[W]hich is well entrenched in its possessor, something that ... “goes all the way down,” unlike a habit such as being a tea-drinker—but the disposition in question, far from being a single track disposition to do honest actions, or even honest actions for certain reasons, is multi-track. It is concerned with many other actions as well, with emotions and emotional reactions, choices, values, desires, perceptions, attitudes, interests, expectations and sensibilities. To possess a virtue is to be a certain sort of person with a certain complex mindset.⁵⁴

- D. *Becoming virtuous is an active process requiring practical wisdom.* Aristotle famously said that “we become just by doing just acts, temperate by doing temperate acts, brave by doing brave acts.”⁵⁵ However, a virtuous person has not merely been habituated to act in a particular way. Rather, she has practical wisdom – the ability to perceive the “morally salient” features of a situation and

47 See Solum, *supra* note 37, at 6; JAMES HECKAMN, GIVING KIDS A FAIR CHANCE 20–22 (2013) (outlining the impact of childhood deprivation on cognitive functioning).

48 Nussbaum, *supra* note 29, at 452.

49 *Id.* at 448. As Nussbaum explains, this helps to insulate Aristotle from the charge of cultural relativism.

50 *Id.* at 449.

51 Roger Crisp, *Introduction*, in ARISTOTLE: NICOMACHEAN ETHICS vii, xvii (Roger Crisp ed., 2014).

52 This is the “standard list” of virtues identified by most contemporary eudaimonistic virtue theorists: Liezl Van Zyl, *Eudaimonistic Virtue Ethics*, in THE ROUTLEDGE COMPANION TO VIRTUE ETHICS 183, 186 (Lorraine Besser-Jones & Michael Slote eds., 2015).

53 Annas, *supra* note 36, at 40.

54 Hursthouse, *supra* note 40.

55 ARISTOTLE, *supra* note 38, at 23.

deliberate correctly about what virtue entails in that context.⁵⁶ Each virtue sits at the mean between excess and deficiency (i.e. courage sits between the vices of fearfulness and rashness; generosity sits between the vices of stinginess and wastefulness).⁵⁷ A person with practical wisdom has developed, through experience, reflection, and deliberation, the ability to strike the right balance.⁵⁸ The exercise of reason is central to this account of virtue, defining which ends are pursued, selecting the appropriate means, and molding affective responses.⁵⁹

- E. *How do we specify the virtues in a particular context?* This paper approaches this process by ascertaining the distinctive features of a particular exercise of choice and then assessing what it means to choose well in that context, identifying those dispositions that reflect success within that sphere and the behavioral biases which need to be overcome.⁶⁰ Dispositions qualify as virtues if they cherish ends that have intrinsic value,⁶¹ contribute to our flourishing and to the flourishing of others,⁶² and accord with our distinctive function as rational beings.⁶³ For example, generosity, which involves giving to “the right people, in the right amounts, at the right time ... [a]nd ... with pleasure, or at least without pain,”⁶⁴ requires the exercise of practical wisdom, involves cherishing an end that has intrinsic value (friendship or social justice), and helps us to flourish by building stronger relationships and communities. One could take a different approach to specifying the virtues without compromising the argument that nudges should promote virtue and discourage vice.
- F. *The fundamental aim of the law should be to promote eudaimonia.* The “end and purpose of a polis is the good life, and the institutions of social life are means to that end.”⁶⁵ Laws should promote that end by enabling humans to “acquire, maintain, and exercise” the virtues.⁶⁶ As Raz argues, “the goal of all political action [is] to enable individuals to pursue valid conceptions of the good and to discourage evil or empty ones.”⁶⁷ Lawmakers should

56 NANCY SHERMAN, *THE FABRIC OF CHARACTER: ARISTOTLE’S THEORY OF VIRTUE*, 3–4, 7 (1991) (explaining that “character is expressed in what one *sees* as much as what one *does*”).

57 Crisp, *supra* note 51, at xvi.

58 Annas, *supra* note 36, at 1.

59 Nussbaum, *supra* note 25, at 180.

60 See Colin Farrelly & Lawrence B. Solum, *An Introduction to Aretaic Theories of Law*, in *VIRTUE JURISPRUDENCE 1*, 7–16 (Colin Farrelly & Lawrence B. Solum eds., 2008) (using a similar process to define the judicial virtues).

61 See Nussbaum, *supra* note 25, at 182–183 (emphasizing that the goods that are valuable to a human life are plural and incommensurable and explaining Aristotle’s view that “the virtues, and their orderly arrangement, represent a set of commitments to cherish all the valuable things”).

62 See Hursthouse, *supra* note 20, at 20 (arguing that the virtues must benefit their possessor).

63 See RONALD SANDLER, *CHARACTER AND ENVIRONMENT: A VIRTUE-ORIENTED APPROACH TO ENVIRONMENTAL ETHICS* 17–18 (2007) (arguing on naturalistic grounds that our rationality makes a significant difference in determining what makes a character trait a virtue).

64 ARISTOTLE, *supra* note 38, at 61.

65 ARISTOTLE, *THE POLITICS*, Book III, chap. ix [1280b].

66 Farrelly & Solum, *supra* note 60, at 2.

67 JOSEPH RAZ, *THE MORALITY OF FREEDOM* 133 (1986).

not remain neutral as to the good life but should actively try to shape preferences so that they are consistent with human flourishing.⁶⁸ They should aim to cultivate the right habits and provide the material conditions that enable people to express the virtues.⁶⁹ They should also seek to act virtuously themselves, avoiding corruption, focusing on the well-being of others, and using practical wisdom to balance the pursuit of virtuous goals (such as deciding whether generosity requires prioritizing one worthy cause or another).⁷⁰

In short, this paper contends that lawmakers should be guided by the following principles: the ultimate end of human life, and politics, is *eudaimonia*; laws should aim to promote *eudaimonia* by cultivating virtue among the populace, with the relevant virtue corresponding to the sphere of choice being regulated; laws should promote good habits and discourage bad habits, and furnish the material conditions enabling people to express the virtues; and most importantly, laws should encourage people to exercise their rational capacity.

So what would a “nudge-world” guided by virtue ethics principles look like in practice?⁷¹ Take the example of retirement savings. Governments have an interest in ensuring that people have retirement savings, given that their capacity to live well in the future depends in part on their material well-being.⁷² Therefore, lawmakers must seriously consider the effectiveness of any policy aimed at increasing savings rates. However, the analysis of consequences would not exhaust the inquiry.⁷³ Governments would also consider whether the policy cultivates the right habits and promotes good moral character. They would acknowledge the behavioral bias that leads people to not save for their retirement – an inability to delay gratification – and seek to craft policies that encourage people to temper their desires. They would favor choice-preserving approaches that respect our rational capacities and assist in developing practical wisdom. In sum, a virtue-minded lawmaker would generally favor policies that encourage people to deliberate about what a good retirement looks like for them, such as requiring people to actively decide whether to enroll in a savings plan, above policies that eliminate choice (like mandates) and non-interventionist approaches that do not challenge bad habits.

Three objections to using virtue ethics to inform policy must be addressed at the outset. The first is the situationist critique, which

⁶⁸ Solum, *supra* note 37, at 11. This paper brackets the question whether this approach is preferable to the neutrality characterizing liberal political theories: see generally Zuckert, *supra* note 34.

⁶⁹ Solum, *supra* note 37, at 5–6, *see also* MARTHA C. NUSSBAUM, *CREATING CAPABILITIES: THE HUMAN DEVELOPMENT APPROACH* (2011) (advocating empowering people with the capabilities to live a good life).

⁷⁰ Daniel C. Russell, *What Virtue Ethics Can Learn From Utilitarianism*, in *THE CAMBRIDGE COMPANION TO UTILITARIANISM* 258, 261–62 (Ben Eggleston & Dale E. Miller eds., 2014).

⁷¹ Waldron, *supra* note 13.

⁷² See generally Sendhil Mullainathan & Eldar Shafir, *Scarcity: Why Having Too Little Means So Much* (2013) for an observation that poverty taxes people’s cognitive capacities such that they are less able to exercise their rational capacities in other spheres.

⁷³ Russell, *supra* note 70, at 258–61.

purportedly calls into question the entire virtue ethics project by challenging the existence of robust, cross-situationally consistent character traits. Drawing on social psychology research, situationists argue that behavior can often, if not always, be attributed to the situation an agent finds herself in rather than her disposition.⁷⁴ For present purposes, it suffices to make two responses to the situationist challenge.⁷⁵ First, virtue is an exacting ideal, requiring actions, emotions, and motivations to meet a demanding standard, and evidence that people fall short of that ideal does not discredit the pursuit of virtue as a guiding principle.⁷⁶ Second, situationist research can enrich virtue ethics by drawing attention to the subtle environmental cues that affect whether people realize their ethical ideals and enabling interventions to be designed that account for those cues.⁷⁷

The second objection is that a virtue-centered theory of legislation has unrealistic aspirations. Laws are crude instruments – they cannot ensure that people do the right thing for the right reasons, or have the appropriate emotional state.⁷⁸ One might question whether nudges, which merely alter the choice architecture within which decisions are made, are capable of cultivating virtue. However, this paper will argue that laws, including choice-preserving approaches like nudges, have the potential to cultivate good character in four ways. First, they can habituate people to act virtuously in a particular context, making it more likely that they could translate that behavior into their lives more generally. Second, they can shape people’s affective responses, prompting positive associations with virtuous behavior and negative associations with vicious behavior. Third, they can aid in the development of practical wisdom, by drawing attention to morally significant features of a situation that might otherwise be neglected. Finally, they can help to furnish the material conditions that may be necessary to express the virtues. Law could never be the sole (or even the predominant) instrument for promoting virtue in a society, but it can steer people in the right direction.

The third objection is that virtue ethics cannot be operationalized. It is true that determining the effects of law on character is an imprecise exercise. We can measure people’s changes in behavior but not whether those changes reflect an altered inner state across the course of their life. However, translating any ethical theory to lawmaking has its difficulties – consequentialists must grapple with the challenge of identifying stable and informed preferences and determining which consequences to prioritize;⁷⁹

⁷⁴ See, e.g., Maria Merritt, John M. Doris & Gilbert Harman, *Character*, in *THE MORAL PSYCHOLOGY HANDBOOK* 356–357 (2010).

⁷⁵ See Gopal Sreenivasan, *The Situationist Critique of Virtue Ethics*, in *THE CAMBRIDGE COMPANION TO VIRTUE ETHICS* 290 (Daniel C. Russell ed., 2013) for a recent review of and response to the situationist critique.

⁷⁶ See Appiah, *supra* note 35, at 48–50.

⁷⁷ *Id.*

⁷⁸ Solum, *supra* note 37, at 10. See also ROBERT GEORGE, *MAKING MEN MORAL: CIVIL LIBERTIES AND PUBLIC MORALITY* 1 (1993) (emphasizing that laws “cannot make men moral”).

⁷⁹ For example, consequentialists must decide whether cost-benefit analysis or well-being analysis better captures the ‘consequences’ of a particular policy: see see John Bronsteen Christopher Buccafusco & Jonathan S. Masur, *Forty-Third Annual Administrative Law Symposium: A Happiness*

deontologists must determine how to resolve apparent conflicts between rights. Virtue ethics directs lawmakers to encourage virtuous actions, motivations, and emotions to the greatest extent possible. When it comes to questions of ethics, any further precision is unattainable and undesirable.

IV. VIRTUE-BASED ANALYSIS OF NUDGING

This section is divided into two parts. The first considers three nudges that have been used to promote sustainable energy consumption: default rules, active choosing, and social norms. The second considers three nudges that have been used to address smoking: factual disclosure, graphic warnings, and precommitment strategies. Each part concludes by explaining how a virtue-based analysis would inform the comparison between these approaches.

A. SUSTAINABLE ENERGY

Governments across the United States are trying to tackle climate change by promoting responsible energy use and accelerating the transition from “gray” to “green” sources of energy.⁸⁰ At the federal level, there is limited political support for market mechanisms like emissions trading schemes and carbon taxes,⁸¹ and the Supreme Court has stalled executive branch attempts to regulate emissions from coal-fired power plants.⁸² Nudges have the potential to achieve environmentally friendly outcomes while avoiding some of these constraints.

1. Default Rules

Default rules are “settings that apply, or outcomes that stick, when individuals do not take active steps to change them ... [they] establish what happens if people do nothing at all.”⁸³ Shifting from “opt in” to “opt out” approaches has had striking outcomes in a number of contexts, including dramatically increasing organ donation rates in Austria⁸⁴ and retirement savings in Denmark.⁸⁵ With respect to sustainable energy, Sunstein and Reisch have observed that switching from a gray to a green default could have environmental outcomes “potentially far larger than the effects of information, education, moral exhortation, and even significant economic

Approach to Cost-Benefit Analysis: Well-Being Analysis vs. Cost-Benefit Analysis, 62 DUKE L. J. 1603 (2013).

⁸⁰ The term “green” is used here as a shorthand for an energy source that emits “lower levels of greenhouse gases and conventional pollutants such as particulate matter and sulfur dioxide” than its “gray” alternative: Cass R. Sunstein & Lucia A. Reisch, *Automatically Green: Behavioral Economics and Environmental Protection*, 38 HARV. ENVTL L. REV. 127, 128 (2014).

⁸¹ Editorial, *Proof that a Price on Carbon Works*, N.Y. TIMES, (Jan. 19, 2016), <https://www.nytimes.com/2016/01/19/opinion/proof-that-a-price-on-carbon-works.html>.

⁸² Lyle Denniston, *Carbon Pollution Controls Put on Hold*, SCOTUSBLOG (Feb. 9, 2016, 6:45 PM), <http://www.scotusblog.com/2016/02/carbon-pollution-controls-put-on-hold/>.

⁸³ Sunstein & Reisch, *supra* note 81, at 131.

⁸⁴ Cass R. Sunstein, *Deciding by Default*, 162 U. PA. L. REV. 1, 4, n.6 (2013).

⁸⁵ Raj Chetty, John N. Friedman, Soren Leth-Petersen, Torben Nielsen & Tore Olsen, *Active vs. Passive Decisions and Crowd-out in Retirement Savings Accounts: Evidence from Denmark* 38 (Nat’L Bureau of Econ. Research, Working Paper No. 18565, 2012), <http://www.nber.org/papers/w18565>.

incentives.”⁸⁶ In Germany, at a time when less than one per cent of the population used green energy, the use of green defaults in two communities drove green usage rates in those communities above ninety percent.⁸⁷

Green defaults may work for a number of different reasons. Sunstein and Reisch have pointed to three possibilities: first, a green default may be perceived as “an implicit suggestion or endorsement” from the choice architect;⁸⁸ second, inertia and procrastination may lead people to stick with the status quo rather than incur the effort of choosing a different energy provider;⁸⁹ and third, green defaults might enlist loss aversion by increasing the salience of the losses incurred by shifting from green energy to gray energy.⁹⁰ Green defaults might also be effective because they trigger guilt about opting out of an environmentally preferable action.⁹¹

2. *Active Choosing*

Active choosing involves requiring or prompting people to make a decision. Instead of relying on inertia to achieve particular outcomes (default rules), active choosing overcomes inertia by ensuring people make (or consider making) a choice.⁹² Active choosing could take a number of different forms in a sustainable energy context. People could be *required* to make a choice between green and gray energy, or merely *asked* to make such a choice.⁹³ Choices could be framed neutrally or consciously encourage a particular energy source.⁹⁴

A recent online experiment conducted by Hedlin and Sunstein has demonstrated the potential of active choosing to promote green energy use.⁹⁵ Their central finding was that active choosing had larger effects in promoting green energy use than green defaults, particularly where green energy was seen as more expensive, a result they hypothesized was due to the “interaction between people’s feelings of guilt and reactance.”⁹⁶ Active choosing was said to trigger guilt, leading people to choose green energy; green defaults triggered resentment, leading some people to opt out.⁹⁷ Despite the limitations of this experimental design,⁹⁸ it is certainly plausible

86 Sunstein & Reisch, *supra* note 81, at 131.

87 *Id.* at 135.

88 *Id.* at 140.

89 *Id.* at 141.

90 *Id.* at 143–44.

91 Aristeidis Theotokis & Emmanouela Manganari, *The Impact of Choice Architecture on Sustainable Consumer Behavior: The Role of Guilt*, 131 J. BUS. ETHICS 423, 424 (2015).

92 Sunstein & Reisch, *supra* note 81, at 151.

93 CASS R. SUNSTEIN, CHOOSING NOT TO CHOOSE: UNDERSTANDING THE VALUE OF CHOICE 87–88 (2015).

94 Sunstein & Reisch, *supra* note 81, at 154 (describing the consciously encouraging approach as “enhanced” active choosing).

95 Simon Hedlin & Cass R. Sunstein, *Does Active Choosing Promote Green Energy Use? Experimental Evidence*, 43 ECOL. L.Q. 107 (2016).

96 *Id.* at 113–14.

97 *Id.* at 112.

98 Most significantly, participants in the experiment were prompted to consider the desirability of a green default, likely limiting the impact of inertia and procrastination and increasing the opt-out rate; see *id.* at 113.

that active choosing could be more effective than default rules in achieving green outcomes in certain contexts.⁹⁹

3. *Social Norms*

Social norms nudging is built on the insight that “the perceived decisions of others can have a big influence on individual behavior and choice.”¹⁰⁰ One reason this might occur is because insufficiently informed individuals accept a belief because it is accepted by others, leading to a “bandwagon or snowballing process” (an “informational cascade”).¹⁰¹ Another reason is that people might act and think in a particular way to “earn social approval and avoid disapproval,” prompting a “reputational cascade.”¹⁰² There are two types of norms which lawmakers might deploy to trigger these cascades – descriptive norms, which provide information about “what is typical or normal,” and injunctive norms, which specify “what constitutes morally approved and disapproved conduct.”¹⁰³

Social norms have been used by choice architects for a range of purposes, including reducing alcohol consumption on university campuses,¹⁰⁴ increasing tax compliance,¹⁰⁵ and encouraging charitable giving.¹⁰⁶ However, its most promising application may be to promote household energy conservation. As part of a large randomized controlled trial, the American energy provider, OPOWER, sent Home Energy Report (HER) letters to residential utility customers across the United States comparing their electricity use to that of their neighbors (a descriptive norm). Those comparisons were accompanied by “smiley face” emoticons for customers who used less energy than their neighbors (an injunctive norm). Reviewing this initiative, Allcott concluded that the average program reduced energy consumption by 2%, an effect equivalent to that of a short-term electricity price increase of 11 to 20%.¹⁰⁷ Significantly, HERs appeared to have had “constant or increasing effects as they [were] repeatedly delivered over the first two years of treatment.”¹⁰⁸ Accordingly,

99 See also Theotokis & Manganari, *supra* note 92, at 432, who found that a forced choosing policy was as effective or more effective than a green default in various settings.

100 CASS R. SUNSTEIN, *SIMPLER: THE FUTURE OF GOVERNMENT* 65 (2013).

101 Timur Kuran & Cass R. Sunstein, *Availability Cascades and Risk Regulation*, 51 STAN. L. REV. 683, 685–86 (1999).

102 *Id.* at 686.

103 Robert B. Cialdini, Raymond R. Reno & Carl A. Kallgren, *A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places*, 58 J. PERSONALITY & SOC. PSYCHOL. 1015, 1015 (1990)

104 H. Wesley Perkins, *Social Norms and the Prevention of Alcohol Misuse in Collegiate Contexts*, 14 J. STUD. ALCOHOL S-164 (2002).

105 HALPERN, *supra* note 6, at 112–14.

106 Bruno Frey & Stephan Meier, *Social Comparisons and Pro-Social Behavior: Testing ‘Conditional Cooperation’ in a Field Experiment*, 94 AM. ECON. REV. 1717 (2004).

107 Hunt Allcott, *Social Norms and Energy Conservation*, 95 J. PUB. ECON. 1082, 1083 (2011). See also Ian Ayres *et al.*, Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage (Nat’l Bureau of Econ. Research, Working Paper No. 15386, 2009), <http://www.nber.org/papers/w15386.pdf>.

108 *Id.* (explaining that this effect was primarily attributed to the use of the descriptive norm rather than the injunctive norm).

social norms nudging has the potential to shift energy consumption at the cost of a letter.¹⁰⁹

4. *Virtue-Based Analysis*

Ethical debates about default rules, active choosing, and the use of social norms have primarily centered on consequentialist and deontological considerations. Issues considered include whether default rules violate autonomy,¹¹⁰ whether required choosing inhibits welfare,¹¹¹ and whether social norms can be used to promote lasting behavior change.¹¹² A virtue-based analysis allows us to take the ethical debate in a different direction, by considering whether these nudges develop character and promote human flourishing.

To answer that question, we must first identify the relevant environmental virtues.¹¹³ As outlined in the model theory, this requires considering the distinctive features of energy consumption and assessing which character traits would cherish the right ends and promote human flourishing. First, we can easily recognize that the reckless use of energy by individuals, when repeated across a population, contributes to the overall problem of climate change. Accordingly, the decision to consume should be influenced by the virtue of benevolence, which involves caring about the wellbeing of others, including future generations.¹¹⁴ Second, energy consumption decisions are shaped by our capacity to regulate our desires. Considering the financial cost of overconsumption, and the environmental implications, these decisions should be guided by the virtue of temperance, which counsels us to consume in proportion to our needs.¹¹⁵ Third, decisions about energy affect our shared environment, which is both an intrinsically valuable end and instrumentally valuable through providing the material conditions that enable us to flourish. Accordingly, we should be guided by the virtue of respect for nature.¹¹⁶ Finally, humans should

109 Social norms have been used successfully to promote other pro-environmental behaviors, such as the purchasing of eco-friendly products: see Christophe Demarque *et al.*, *Nudging Sustainable Consumption: The Use of Descriptive Norms to Promote a Minority Behavior in a Realistic Online Shopping Environment*, 43 J. ENV. PSYCHOL. 166 (2015).

110 See Rebonato, *supra* note 7, at 370.

111 See Sunstein, *supra* note 94, at 113–15.

112 See Allcott, *supra* note 109, at 1083 (discussing concern over the durability of treatment effects in non-price interventions).

113 Ronald Sandler describes four different strategies which have been used to specify environmental virtues: arguing by extension from standard interpersonal virtues; appealing to agent benefit; arguing from considerations of human excellence; and studying the character traits of exemplary individuals: Ronald Sandler, *Introduction*, in ENVIRONMENTAL VIRTUE ETHICS 1, 4–5 (Ronald Sandler & Philip Cafaro eds., 2005). This paper uses the first three strategies, consistent with the model theory outlined above.

114 See Geoffrey Frasz, *Benevolence as an Environmental Virtue*, in ENVIRONMENTAL VIRTUE ETHICS 121, 124 (Ronald Sandler & Philip Cafaro eds., 2005) (arguing that benevolence “improves the life of the benevolent person ... directly, through the satisfaction of acting morally, and indirectly, through helping to create a better society for all.”)

115 See Peter Wenz, *Synergistic Environmental Virtues*, in ENVIRONMENTAL VIRTUE ETHICS 197, 208 (Ronald Sandler & Philip Cafaro eds., 2005) (arguing that rational consumers practice temperance and avoid overconsumption).

116 See Sandler, *supra* note 63, at 6 (arguing that the virtue of respect for nature is justified by the intrinsic worth of the environment).

consume energy in a distinctively human way, exercising their rational, deliberative capacities to make mindful energy choices.

Viewed through this lens, green defaults are problematic, at least when compared to active choosing. Default rules work primarily because of inertia and procrastination. They achieve environmentally friendly outcomes without spurring us to exercise our rational capacities to make sustainable choices. Green defaults may lead us to consume energy in ways consistent with a benevolent and temperate agent who respects nature. However, their success does not depend on us having any particular attitudes or dispositions, or cherishing any particular ends. By not requiring the exercise of choice, they are likely to prompt short-term, context-specific improvements in behavior without fostering a meaningful engagement with the environmental implications of our consumption.

Default rules do have some potential advantages. First, by making energy consumption decisions automatic, green defaults can free up cognitive bandwidth so that we could exercise our capacities for rational thought in other spheres.¹¹⁷ Lawmakers may do better by encouraging ethical deliberation in contexts other than energy consumption, considering the complexity of the science and the degree of attention required. Second, to the extent that green defaults are seen as an implicit recommendation from lawmakers, they could provide people with information about what sustainable consumption requires, making it easier for them to act virtuously. If lawmakers are trusted, green defaults might also signal that green energy is desirable, prompting the right emotional associations. Third, if green defaults trigger guilt, they could help people recognize the moral significance of their decision to opt out. However, the latter two advantages depend on people being aware that they have been defaulted into a green energy plan, which is not necessarily the case.

Active choosing has an important advantage over default rules in this context. It requires people to exercise their rational capacities to choose between different energy plans, thus cultivating practical wisdom by making us more mindful of our energy consumption. Active choosing also better caters for heterogeneity, spurring us to make energy choices that reflect our own conception of human flourishing. However, here as elsewhere, “active choosing is not a panacea.”¹¹⁸ Requiring people to choose between different energy plans does not ensure that they choose the sustainable option, and, even if they do, it does not ensure that their motives for that choice and associated emotions comport with the virtues of benevolence, respect for nature, and temperance. Accordingly, the best approach may be a form of “enhanced” active choosing,¹¹⁹ in which policymakers make it clear that a particular choice is environmentally preferable and seek to trigger positive affective reactions to green options (or guilt when choosing gray options). An approach of this type enriches

117 See Cass R. Sunstein, *Autonomy by Default*, 16 AM.J.BIOETHICS 1, 1 (2016) (arguing that “to the extent that default rules give people the freedom to focus on their most pressing concerns, and thus eliminate a kind of ‘bandwidth tax,’ they increase autonomy as well.”)

118 Sunstein & Reisch, *supra* note 81, at 154.

119 *Id.*

the deliberation between different energy plans, responding to Waldron's concern that nudges do not help one become a "better chooser."¹²⁰

The appropriateness of social norms nudging depends on the nature of the social norm being deployed. A descriptive norm comparing one's energy consumption to that of one's neighbors is potentially useful, helping people become mindful of their energy consumption and giving them a rough benchmark against which they can judge whether their consumption is temperate.¹²¹ An injunctive norm encouraging people for good behavior (such as a smiley face) could help to habituate virtuous consumption.

However, social norms nudging risks prompting people to tackle their energy consumption for the wrong reasons. Further research is needed to determine whether the norms deployed in the OPOWER study worked primarily because people felt a sense of competitiveness with their neighbors, wanted to save money, wanted to avoid the reputational consequences of over-consumption, or were triggered to consider their environmental footprint. A virtue-based analysis would support using social norms to emphasize people's contribution to the goal of environmental sustainability, rather than focusing on the money people saved relative to their neighbors. Although that approach may be less effective in spurring behavior change in the short-term, it may be necessary to accept some reduction in effectiveness in exchange for cultivating the appropriate environmental ethic. Further research is also needed to determine whether nudges of this type promote a broader environmental consciousness or merely prompt environmentally friendly behavior in a specific context.

5. Conclusion

Choice architects cannot guarantee that people who shift their energy consumption in response to default rules, active choosing, or social norms have done so for the right reasons. However, a virtue ethics perspective still provides some important guidance for lawmakers, by counseling them to frame nudges to cultivate the appropriate environmental ethic. The ethic governing energy consumption requires people to respect nature, recognize the impact of their consumption on others, and moderate their desires. Green defaults are the least promising nudging technique in this regard, although they do potentially have the expressive effect of emphasizing the desirability of green energy. Active choosing is the most attractive option, cultivating practical intelligence by encouraging mindfulness about energy consumption. However, its success in cultivating the virtues may hinge on policymakers framing choices to encourage green consumption, triggering the appropriate attitudes towards environmentally friendly options. Social norms may also play a role in promoting the environmental virtues.

¹²⁰ Waldron, *supra* note 13.

¹²¹ Of course, this would not hold if most people in a neighborhood were consuming well beyond their means. In those circumstances, policymakers may need to inform people where their consumption compares to an objectively acceptable level.

However, it is essential that the norms deployed target the right character traits.

B. SMOKING

Tobacco use is the single largest preventable cause of death and disease in the United States.¹²² In 2014, an estimated forty million American adults were cigarette smokers, with the prevalence of smoking highest among those below the poverty line.¹²³ The World Health Organization has concluded that raising taxes on tobacco is the most effective way to curb smoking.¹²⁴ However, recent attempts to increase federal cigarette taxes have been repeatedly stymied by Congress.¹²⁵ Nudges are attractive alternatives.

1. *Factual Disclosure*

Factual disclosure of the risks of smoking can be accomplished in a number of different ways and through a range of different mediums. This paper uses the term ‘factual disclosure’ to refer to the textual presentation of the risks of smoking on cigarette packets. The following warning labels are presently mandated in the United States:

SURGEON GENERAL’S WARNING: Smoking Causes Lung Cancer, Heart Disease, Emphysema, and May Complicate Pregnancy

SURGEON GENERAL’S WARNING: Quitting Smoking Now Greatly Reduces Serious Risks to Your Health

SURGEON GENERAL’S WARNING: Smoking by Pregnant Women May Result in Fetal Injury, Premature Birth, and Low Birth Weight

SURGEON GENERAL’S WARNING: Cigarette Smoke Contains Carbon Monoxide¹²⁶

There is no neutral way to present the ‘facts’ about smoking. The framing of the objective risks of smoking impacts how they are subjectively perceived. A message that emphasizes the potential gains of quitting smoking (i.e. ‘You will live longer if you quit smoking’) may be more effective in discouraging smoking than a message that emphasizes the potential losses of not quitting smoking (i.e. ‘You will die sooner if you do

¹²² Centers for Disease Control and Prevention, *Current Cigarette Smoking Among U.S. Adults Aged 18 Years and Older*, (Dec. 14, 2015), <http://www.cdc.gov/tobacco/campaign/tips/resources/data/cigarette-smoking-in-united-states.html>.

¹²³ *Id.*

¹²⁴ *Raising Tax on Tobacco: What You Need to Know*, WORLD HEALTH ORG. (2014), <http://www.who.int/campaigns/no-tobacco-day/2014/brochure/en/>.

¹²⁵ Congress did not act on President Obama’s proposal for a 94-cents-per-pack increase in cigarette taxes in his last four budgets. President Obama was able to pass an increase in the federal cigarette tax from 39 cents to \$1.01 in 2009.

¹²⁶ 15 U.S.C. § 1333 (2012). As discussed below, new graphic warnings (comprising revised textual warnings and graphic images) have been authorized by Congress but have not yet been implemented.

not quit smoking’).¹²⁷ A message that emphasizes less well-known risks of smoking (i.e. heart disease) may be more informative than a message that emphasizes more well-known risks (i.e. lung disease).¹²⁸ Regulators may frame these risks differently depending on whether the primary purpose of disclosure is to inform people of the adverse consequences of smoking or to influence them not to smoke.¹²⁹

For present purposes, it is sufficient to bracket those complexities and focus on disclosure that is intended simply to inform people of the health risks of smoking.¹³⁰ Factual disclosure, in this sense, does not steer people in a particular direction but rather allows them to “decide for themselves whether the costs of smoking outweigh the benefits.”¹³¹ When people have accurate information about the risks of smoking, they are expected to be in a better position to make “careful, deliberate decisions.”¹³²

2. *Graphic Warnings*

Graphic warnings seek to discourage people from smoking by provoking “visceral negative reactions” to cigarettes.¹³³ In 2009, Congress enacted the Family Smoking Prevention and Tobacco Control Act,¹³⁴ which directed the Food and Drug Administration (“FDA”) to issue regulations requiring that all cigarette packages display “color graphics depicting the negative health consequences of smoking.”¹³⁵ The FDA selected a set of nine graphic images, including depictions of diseased lungs, rotting teeth, and children crying. They projected that this regulation would have significant health benefits, estimating that it would reduce the number of smokers by 213,000 in 2013, with smaller additional reductions through 2031.¹³⁶ However, before these graphic warnings could be rolled out, a

127 Benjamin A. Toll et al., *Comparing Gain- and Loss-Framed Messages for Smoking Cessation with Sustained-Release Bupropion: A Randomized Controlled Trial*, 21 *PSYCHOL. ADDICT. BEHAV.* 534, 535, 542 (2007) (drawing on prospect theory).

128 See Ellen Peters, *A Heuristics Approach to Understanding Cancer Risk Perception: Contributions from Judgment and Decision-Making Research*, 31 *ANN. BEHAV. MED.* 45, 47 (2006) (suggesting that the representativeness heuristic leads smokers to overestimate how likely smokers are to contract lung cancer). Even trivial factors like the choice of font in warning labels may affect risk perception: see Christine Jolls, *Product Warnings, Debiasing, and Free Speech: The Case of Tobacco Regulation*, 169 *J. INST. & THEORETICAL. ECON.* 53, 57 (2013).

129 See Memorandum from Cass R. Sunstein, Adm’r, Office of Info. & Regulatory Affairs, to the Heads of Exec. Dep’ts & Agencies, *Disclosure and Simplification as Regulatory Tools* (June 18, 2010), [available at https://www.whitehouse.gov/sites/default/files/omb/assets/inforeg/disclosure_principles.pdf](https://www.whitehouse.gov/sites/default/files/omb/assets/inforeg/disclosure_principles.pdf).

130 The current US warning labels typify this approach. Their purpose was to “enable individuals to make informed decisions about smoking.” Comprehensive Smoking Education Act, Pub. L. No. 98-474, § 2, 98 Stat. 2200, 2200 (1984) (codified at 15 U.S.C. § 1331).

131 Ryan Bubb, *TMI? Why the Optimal Architecture of Disclosure Remains TBD*, 113 *MICH. L. REV.* 1021, 1028 (2015). Of course, the mere act of referring to health risks might steer people in a particular direction by making those risks more salient.

132 *Id.* at 1025.

133 Sunstein, *supra* note 100, at 133.

134 Pub. L. No. 111-31, 123 Stat. 1776 (2009) (codified in various sections of 15 U.S.C. and 21 U.S.C.). The Act also required cigarette packages to bear one of nine new textual warning statements: see Federal Food, Drug, and Cosmetic Act, 21 U.S.C. § 201 (2012). The discussion of graphic warnings that follows focuses on the images rather than the revised text.

135 15 U.S.C. § 1333 (2012).

136 Food and Drug Administration Required Warnings for Cigarette Packages and Advertisements, 21 C.F.R. § 1141 (2011). Subsequently, researchers have estimated that had the

federal appeals court ruled that requiring sellers to display them violated the First Amendment to the U.S. Constitution.¹³⁷ The FDA has announced its intention to issue a new graphic warning rule, but has not yet done so.¹³⁸

Graphic warnings are different from factual disclosure in that they are not intended to be educative.¹³⁹ As Sunstein explains, they work by “directly target[ing] System 1, trying to make the health consequences of smoking salient and to enlist the affect heuristic in the direction of reduced smoking.”¹⁴⁰ They aim to provoke “strong, hostile emotional reactions” to cigarettes by “associating smoking with illness, distress, pain, and death.”¹⁴¹ In short, graphic warnings seek to discourage smoking by triggering hostile emotional reactions; factual disclosure seeks to ensure that smokers are making an informed choice.

3. *Precommitment Strategies*

Precommitment devices have been described as “any means through which consumers impose constraints on their future behavior.”¹⁴² Precommitment works when people are aware of their biases and are willing to commit to strategies to counteract them.¹⁴³ By confining their future selves, people are more likely to be able to overcome present bias and unrealistic optimism.¹⁴⁴ Elster has described this as a form of “rationality over time.”¹⁴⁵ A useful illustration is provided by the website stickK.com, which allows people to put up a certain amount of money that they forfeit unless they achieve a specified goal.¹⁴⁶ Governments in Norway, Canada, and Australia have demonstrated the policy application of this approach, introducing pre-commitment systems for electronic gaming machines.¹⁴⁷

Halpern et al have demonstrated the potential for precommitment to spur smokers to quit smoking.¹⁴⁸ They offered CVS Caremark employees

graphic warnings been implemented in 2012, there would have been a reduction in the number of adult smokers of between 5.3 and 8.6 million in 2013; Jidong Huang et al., *Cigarette Graphic Warning Labels and Smoking Prevalence in Canada: A Critical Examination and Reformulation of the FDA Regulatory Impact Analysis*, 23 TOBACCO CONTROL (Supp. 1) i7 (2014).

137 R.J. Reynolds Tobacco Co. v. FDA, 696 F.3d 1205 (D.C. Cir. 2012).

138 In *Discount Tobacco City and Lottery v. United States*, 674 F.3d 509 (6th Cir. 2012), the United States Court of Appeals for the Sixth Circuit held that the FDA did have the authority to promulgate a graphic warning rule. Those conflicting rulings have the effect that the FDA does have the authority to require graphic warnings, but needs to select different warnings to those chosen previously.

139 However, as is discussed below, graphic warnings may have the incidental effect of correcting misperceptions about the risks of smoking.

140 Sunstein, *supra* note 100, at 132.

141 *Id.* at 133.

142 Peggy J. Liu et al., *Using Behavioral Economics to Design More Effective Food Policies to Address Obesity*, 36 APPL. ECON. PERSP. POL'Y. 6, 13 (2014).

143 Cass R. Sunstein, *Nudging Smokers*, 372 N. ENGL. J. MED. 2150, 2151 (2015).

144 *Id.*

145 JON ELSTER, *ULYSSES UNBOUND: STUDIES IN RATIONALITY, PRECOMMITMENT, AND CONSTRAINTS* 5 (Cambridge Univ. Press 2000).

146 See generally IAN AYRES, *CARROTS AND STICKS: UNLOCK THE POWER OF INCENTIVES TO GET THINGS DONE* xv–xvi (Bantam Books 2010).

147 See generally Zeb Kurth-Nelson & A. David Redish, *Don't Let Me Do That! – Models of Precommitment*, 6 FRONT NEUROSCI. 138 (2012).

148 Scott D. Halpern et al., *Randomized Trial of Four Financial-Incentive Programs for Smoking Cessation*, 372 N. ENGL. J. MED. 2108 (2015).

and their friends and relatives the opportunity to participate in two different smoking cessation programs. The first offered participants a reward of \$800 for smoking cessation (reward program). The second offered participants \$650 in reward payments for successful completion, but also required them to contribute a refundable deposit of \$150 (deposit program). Both programs used precommitment to make it easier for smokers to quit smoking, by denying rewards (reward program) or imposing costs (deposit program) on those who failed. However, they had very different results. The reward program was much more successful in encouraging people to sign up to a smoking cessation program, but the deposit program was significantly more effective in leading those who selected it to stop smoking.¹⁴⁹ That latter finding can be primarily attributed to the power of loss aversion. As people are “typically more motivated to avoid losses than to seek gains,” participants that put their own money at risk were more likely to change their behavior.¹⁵⁰

4. *Virtue-Based Analysis*

Each of the nudges surveyed above has been challenged on consequentialist grounds. There are extensive debates about whether (and to what extent) graphic warnings are more effective than factual disclosure in reducing smoking rates,¹⁵¹ whether graphic warnings impose an emotional tax that outweighs their welfare benefits,¹⁵² and whether precommitment strategies undermine learning.¹⁵³ Graphic warnings, in particular, have been challenged on deontological grounds for manipulating people to refrain from smoking.¹⁵⁴

A virtue-based approach leads policymakers to consider another important issue: which of these nudges cultivates the virtues that ought to be associated with the decision to smoke? To identify these virtues, we need to determine what it means to choose well within this sphere of human experience. Two distinctive features of smoking inform this analysis. First, the decision to smoke is driven primarily by the desire for pleasure. Desiring pleasure is natural for human beings.¹⁵⁵ However, we can easily recognize that humans often need to moderate their desires in order to flourish. A person who consistently eats, drinks, or smokes to excess risks compromising her health and undermining her life prospects. Addiction degrades the distinctive human capacity to make rational

¹⁴⁹ *Id.* at 2114. As reward-based programs were significantly more commonly accepted than deposit-based programs, they lead to higher rates of sustained abstinence from smoking overall. *Id.* at 2108.

¹⁵⁰ *Id.* at 2109. On the power of loss aversion, see Ayres, *supra* note 146, at 45–71.

¹⁵¹ For an empirical analysis of the effectiveness of various health warnings, see David Hammond et al., *Communicating Risk to Smokers: The Impact of Health Warnings on Cigarette Packages*, 32 AM. J. PREV. MED. 202 (2007).

¹⁵² See, e.g., Edward L. Glaeser, *Paternalism and Psychology*, 73 U. CHI. L. REV. 133, 135 (2006) (arguing that soft paternalistic policies, like those that stigmatize smoking, essentially impose a “psychic tax” that provides no revenues).

¹⁵³ See Bailey Kuklin, *Self-Paternalism in the Marketplace*, 60 U. CIN. L. REV. 649, 666 (1992).

¹⁵⁴ See generally Sunstein, *supra* note 14.

¹⁵⁵ On this point, virtue ethicists do not differ from utilitarians.

choices, even if smokers make a rational decision to take up smoking.¹⁵⁶ Accordingly, the virtue of temperance should regulate the desire to smoke – those who do choose to smoke should do so in moderation. Second, smoking may cause significant health risks to others. Accordingly, the decision to smoke should be governed by benevolence, the virtue that “disposes us to care about other people’s flourishing.”¹⁵⁷ A benevolent smoker would avoid smoking in contexts where it would harm others, such as during pregnancy.

Policymakers should regulate smoking in ways that promote human flourishing. In some situations, this may require mandates – banning children from smoking recognizes that they have not developed the rational capacities to decide whether to smoke; banning smoking in public places may be necessary if people refuse to benevolently regulate their cigarette consumption. However, policymakers should be predisposed to giving people the choice to smoke, while encouraging smokers to do so in a manner consistent with the virtues of temperance and benevolence. In evaluating all interventions in this sphere, there is a pervasive question of effectiveness – as addiction undermines the capacity of smokers to flourish, the effectiveness of policies to forestall addiction is an important consideration.¹⁵⁸

On the surface, factual disclosure is an attractive option. It respects our rational capacities by seeking to educate us about risks, enabling reasoned deliberation about the decision to smoke. However, there is a substantial body of evidence suggesting that factual text-based warnings are ineffective as informational tools.¹⁵⁹ As smokers become acclimatized to different factual warnings, they lack emotional salience and have a limited effect on decision-making.¹⁶⁰ Despite a long history of factual disclosure in the U.S., there remain significant misperceptions among smokers about the risks of smoking.¹⁶¹ Accordingly, factual disclosure does not necessarily provide people with the information to temper their desires or regulate their smoking in the interests of others, rendering it problematic as a tool to help people express the virtues. By not forestalling addiction, factual disclosure may also be criticized for failing to preserve our distinctively human rational capabilities.

Graphic warnings, on the other hand, appear to be problematic from a neo-Aristotelian perspective. They seek to sidestep our deliberative capacities by triggering strong emotional responses to cigarettes. However,

156 Compare W. KIP VISCUSI, *SMOKING: MAKING THE RISKY DECISION* (1992). Tobacco addictions may spur a form of “tunneling,” leading smokers to neglect other considerations: see the discussion of tunneling in Mullainathan & Shafir, *supra* note 72, at 29–30.

157 Heidi Li Feldman, *Prudence, Benevolence, and Negligence: Virtue Ethics and Tort Law*, 74 CHI.-KENT L. REV. 1431, 1439-1440 (2000).

158 One option would be to ban smoking, if that was shown to be the most effective way to forestall addiction. Regardless of the political viability of this proposal, it is not desirable, as completely eliminating the choice to smoke impedes the development of practical wisdom.

159 See COMM. ON Preventing Nicotine Addiction IN Children & Youths, Inst. OF Med., *Growing Up Tobacco Free: Preventing Nicotine Addiction IN Children AND Youths* 240 (Barbara S. Lynch & Richard J. Bonnie eds., 1994).

160 Sunstein, *supra* note 100, at 130.

161 Jolls, *supra* note 128, at 58–59.

graphic warnings do have a number of advantages. First, as has been emphasized, virtue requires motives and emotions to be aligned with actions. By triggering hostile affective responses to smoking, graphic warnings may help to bring actions and attitudes towards smoking into conformity for those who want to quit but cannot, enabling the exercise of temperance. Second, they could help draw our attention to salient features of a situation, including the impact of smoking on non-smokers, which is essential to forging a benevolent disposition. Third, considering the harms of addiction to our rational capacities, graphic warnings may be essential to enabling us to express the virtues. By dissuading us from smoking, they make it less likely that we will face the health consequences of smoking and help to preserve our rational decision-making capacities. Fourth, graphic warnings may in fact be more effective in educating people about the risks of smoking than factual disclosure, even though they are not necessarily intended to be educative. As Jolls has shown, graphic warnings can help to correct factual misperceptions about the risks of smoking.¹⁶² Debiasing people through graphic warnings could have the effect of enriching people's exercise of their deliberative capacities.

Graphic warnings do risk overshooting the mark in some respects. Encouraging people to view the decision to smoke with a temperate and benevolent disposition does not necessarily require revulsion at the prospect of smoking per se. A preferable approach would be to use graphic warnings to viscerally illustrate some of the harmful effects of smoking (such as through using images of non-smokers affected by smoking), or to emphasize the risks of intemperate consumption, rather than implying that smoking always merits a hostile response. However, graphic warnings may be too crude a tool to communicate those types of subtle messages.

Precommitment strategies also have favorable characteristics. They empower our deliberative capacities by requiring us to bind ourselves to future courses of action. They cultivate temperance by spurring us to moderate our future desires. In this context, significantly, they only apply to those who have opted-in to a smoking cessation program, respecting the outcome of deliberative processes.

However, precommitment strategies also have a number of limitations. First, in the forms considered in this paper, they rely on loss aversion or the prospect of reward to spur smoking cessation, rather than intrinsic motivations to quit. Accordingly, they may lead people to do the right thing for the wrong reason. Second, programs of this type are likely to only promote discrete improvements in specific contexts, potentially curbing smoking without helping people temper their desires in other contexts.¹⁶³ However, getting people to quit smoking would be a significant achievement that enables a broader range of capabilities.¹⁶⁴ If

¹⁶² *Id.* at 54. However, compare Bubb, *supra* note 131, at 1030–1034 (challenging the view that graphic warnings are an effective debiasing tool).

¹⁶³ See Ayres, *supra* note 146, at 156 (pointing to evidence suggesting that “self-control is a limited resource that can be depleted if it is overused”).

¹⁶⁴ *Id.* at 154 (observing that people who successfully quit smoking often put on a few pounds through eating more, but still add years to their life expectancy).

precommitment strategies could be shown to be effective in curbing smoking at scale and at a low-cost, there is a strong case to be made for them.

5. *Conclusion*

A virtue-based analysis generates a number of important insights for policymakers considering nudging smokers. Factual disclosure has the benefit of respecting our deliberative capacities. However, it fails to live up to its promise of informing people of the risks associated with smoking, and accordingly does not enable people to express the virtues. Graphic warnings appear to be superior as informational tools. They also have the advantage of drawing people's attention to salient features of the decision to smoke and triggering appropriate emotional reactions. However, they need to be carefully calibrated to ensure that they actually promote the virtues of temperance and benevolence, rather than inspire an unthinking revulsion towards smoking. Precommitment strategies are a promising attempt to empower our deliberative capacities. Although they may not be able to prompt broader dispositional improvements, their effectiveness in helping addicts retain their deliberative faculties weighs heavily in their favour.

V. LESSONS FOR LAWMAKERS

This paper has argued that virtue ethics should guide the decision to nudge. A virtue-based approach would lead lawmakers to consider issues of virtue, character, and human flourishing that are neglected in current debates about nudging. The model theory sketched in this paper is broad enough to encompass consideration of the consequences of a policy, as consequences will always be relevant to the question whether a policy promotes human flourishing. It can also accommodate deontological concerns, as policies which violate autonomy or dignity would likely conflict with virtue ethics' emphasis on treating people as rational beings. However, it offers a fresh perspective because it does not treat these consequences or rules as dispositive. Instead, it prompts direct consideration of whether these policies enable people to express the virtues. It requires lawmakers to look beyond individual preferences and assess whether policies truly help people flourish.

How can this approach be applied in practice? This paper has offered two illustrations of a virtue-based approach to nudging: the first concerning nudges used to promote sustainable energy consumption; the second concerning nudges used to address smoking. It has argued that all of the tools surveyed have advantages and disadvantages from a virtue ethics perspective; what matters is the framing of these interventions and their intended purpose. However, there are some important broader lessons for choice architects. First, nudges should generally aim to promote rational deliberation, accounting for what is distinctive about human functioning. Accordingly, System 2 nudges like active choosing, factual disclosure, and precommitment, which prompt us to exercise our rational capacities, are

generally preferable to System 1 nudges like default rules and graphic warnings. However, in those circumstances where rational deliberation becomes difficult, such as where nicotine addiction impairs decision-making capacities, System 1 nudges such as graphic warnings may be preferable.

Second, lawmakers should generally favor nudges that play an educative function, helping people express the virtues by enabling them to choose with greater knowledge or understanding. For example, an “enhanced” form of active choosing could be used to promote informed decision-making. That educative function is not always best performed by nudges that are intended to be educative, like factual disclosure. Nudges like graphic warnings, which are not designed to be educative, may in practice work more effectively in providing people with relevant information and drawing attention to the morally salient features of a choice. Accordingly, lawmakers should not be too rigid in assessing the educative potential of nudges, and should test the effectiveness of various approaches in increasing knowledge or understanding.

Finally, lawmakers should not confine their attention to the impact of a policy on behavior, but should also consider whether it promotes the right attitudes and motives. Default rules, for example, are problematic because their success depends on inertia rather than a change in attitudes. Precommitment strategies risk failing to cultivate intrinsic motivations to change behavior, potentially spurring short-term improvements in behavior that are easily reversed. The use of social norms may not be appropriate if it merely encourages people to compete with their neighbors rather than cultivating an appropriate environmental ethic. Although none of these considerations may be dispositive, lawmakers must take them into account in order to fully account for the implications of a policy.