DOI: 10.1007/s10272-018-0713-z

End of previous Forum article

Daniel M. Hausman

Nudging and Other Ways of Steering Choices

One of the most discussed contemporary ideas in policy making is Richard Thaler and Cass Sunstein's proposal to influence people's choices without coercing them by improving the "architecture" of their choices. For example, if it is impossible to get cash from an ATM without first removing one's bank card, then even the careless and forgetful will not leave their cards in the machine. By structuring choices wisely, governments can get people to make better choices without limiting their freedom.

In their book, *Nudge*, Thaler and Sunstein maintain that a nudge is "any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives".[1] In more recent work, Sunstein elaborates on this characterization of nudges as follows:

> Nudges are interventions that steer people in particular directions but that also allow them to go their own way. 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 (including disincentives). 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. …Some nudges work because they inform people; other nudges work because they make

Daniel M. Hausman, University of Wisconsin-Madison, USA.

1 R. Thaler, C. Sunstein: Nudge: Improving Decisions About Health, Wealth, and Happiness, New Haven 2008, Yale University Press, p. 6.

certain choices easier; still other nudges work because of the power of inertia and procrastination.[2]

Nudging is any way of influencing choices that preserves freedom of choice. So defined, nudges could include, at one extreme, slipping some LSD into someone's drink or, at the opposite extreme, leaving Gideons Bibles in hotel rooms.

Lumping together such fundamentally different ways of shaping choices is unhelpful. Apart from coercing people or providing "significant material incentives", one can "steer" people in many different ways:

- *Encouraging or discouraging*: Providing weak or non-material incentives
- *Informing*: Providing relevant information
- *Activating or inciting*: Stimulating emotions to motivate individuals
- *Training or educating*: Influencing their deliberative capacities[3]
- *Deceiving*: Providing false information
- *Brainwashing*: Influencing someone's choice through means such as subliminal images, drugs or hypnotism
- *Nudging* (in a narrow sense): Changing the choice circumstances to neutralize or to exploit deliberative foibles.

Despite some overlap, these methods of influencing behavior differ significantly from one another. With the possible exception of deceit, Thaler and Sunstein would call all of these "nudges". Some of these methods limit freedom of choice. I shall refer to all of these ways of influencing people's choices in a broad sense as "steering", in order to distinguish it from the narrow sense that I shall be concerned with. Although it risks confusion, I shall use the term "nudge" to refer only to the last method of steering people toward some action.

I retain the term "nudge" because the narrow construal of nudges as shaping the circumstances of choice fits the central examples that Thaler and Sunstein give. What is at issue is the importance of distinguishing among the many ways to influence choices, not how one uses the word "nudge." Rather than distinguishing nudging from informing or training, I could distinguish choice-structuring nudges from informational or educational nudges. What makes Thaler and Sunstein's proposal original and important is not the recognition that it is possible to persuade people without coercing them. Long before Mark Antony's funeral oration, leaders were well aware that they could guide people by informing them, encouraging them or stirring up their emotions. What made Thaler and Sunstein's *Nudge* noteworthy was its recognition that the flaws behavioral economics identified in human decision-making constitute both problems and opportunities for public policy. Thaler and Sunstein's contribution rests on what they have to say about nudges in the narrow sense I am employing. In placing healthy foods prominently in a cafeteria line, Thaler and Sunstein's "choice architect", Caroline, is not encouraging or discouraging, informing, activating, training, deceiving or brainwashing individuals. She is instead nudging them. Setting the default in a pension plan nudges employees rather than informing, activating, training, deceiving or brainwashing them.

This taxonomic discussion shows its worth when one turns to the question that occupies Sunstein in his 2015 essay. He asks whether steering people is ethical and, more specifically, whether it promotes or undermines welfare, autonomy and dignity.[4] These questions are not well posed, because their answers vary depending on which means of steering behavior are employed. Providing information does not threaten autonomy or dignity. Brainwashing and deceiving do.

Sunstein is aware that whether steering is ethical depends on how people are to be steered. For example, he writes that whether "nudges intrude on autonomy...depends on what kind of nudge is involved".[5] But Sunstein then relies on the heterogeneity of steering to evade criticisms of nudging in the narrow sense. Consider, for example, Sunstein's response to Sarah Conly's remarks that "(r)ather than regarding people as generally capable of making good choices, we outmaneuver them by appealing to their irrationality, just in more fruitful ways."[6] Sunstein replies, "But she is making a strong charge, one that is not fairly leveled against most kinds of nudges. Recall that many

2 C. Sunstein: The Ethics of Nudging, in: Yale Journal on Regulation, Vol. 32, No. 2, 2015, pp. 413-450, here p. 417.

3 Grüne-Yanoff and Hertwig call this "boosting" rather than training; see T. Grüne-Yanoff, R. Hertwig: Nudge Versus Boost: How Coherent are Policy and Theory?, in: Minds and Machines, Vol. 26, No. 1-2, 2016, pp. 149-183. I prefer "training", because it avoids implying that changes induced in people's deliberation are necessarily positive.

4 C. Sunstein, op. cit., p. 413.

5 Ibid., p. 437.

6 S. Conly: Against Autonomy: Justifying Coercive Paternalism, Cambridge 2012, Cambridge University Press, p. 30.

nudges are educative."[7] The fact that informing people is generally unproblematic, however, tells us nothing about the acceptability of other ways of influencing people, including neutralizing or exploiting deliberative foibles.

No one doubts that informing people or educating them to be better deliberators is not in itself ethically objectionable, although these methods, like any others, can be employed toward malicious ends. Other ways of influencing behavior, such as brainwashing individuals or deceiving them, raise red flags, regardless of the objectives that they aim to achieve. Should policies that influence choices by neutralizing or exploiting deliberative foibles be of ethical concern? My concern is with the appraisal of this method of influencing behavior, not with the objectives to which nudging may be instrumental. The fact that it would be unacceptable to make voting for incumbents the default option says nothing about whether there is anything ethically problematic about nudging itself.[8]

In the case of paternalistic nudges (and nudges need not be paternalistic), Thaler and Sunstein maintain that for nudges to be ethically acceptable, the individuals who are nudged should agree that the actions they are nudged into choosing are better for them than the alternatives. This necessary condition is problematic. First, as Sunstein notes, individuals may retrospectively approve of nudges that steer them toward opposite actions.[9] Second, asking people whether they approve of being nudged poses practical problems. There is also a serious worry that the flaws in the ways in which people deliberate undermine the claims of preferences to determine whether paternalistic nudges succeed in making people better off. If people are not good judges of what to do, why should economists accept their judgment of whether a nudge is good for them? Nudges that aim to benefit people and fail to do so are subject to criticism for their failure. But if policy makers accept the findings that make paternalistic nudging feasible, how are they to determine whether it is desirable?

The choice architect who seeks to nudge people takes them as they are. Rather than seeking to improve their deliberative capacities, the choice architect seeks to structure some choice that people face so as to either neutralize the effects of their deliberative flaws or to harness those flaws to get individuals to choose an alternative that the architect judges to be better. Provided that the objective passes ethical muster, is this method of influencing choices ethically acceptable?

There are two main arguments in defense of the ethical acceptability of nudging. First, Thaler and Sunstein argue that steering people is unavoidable and hence that it must be permissible. However, even if steering is unavoidable, nudging is not. If a despicable politician wants to stir up support for anti-immigrant policies, inciting fear and resentment is an obvious alternative to nudging. Other alternatives to nudging, such as educating people, may not even count as steering them, because it does not influence them "in a particular direction". Nudging is not inevitable.

A second argument in defense of nudging maintains that when it is advisable to influence people's choices – and unproblematic methods such as informing them, training them or encouraging them are ineffective – then nudging is better than coercing people, brainwashing them or stirring up their emotions. If it is important to influence choices, nudging may be the best way to do so. The objections to nudging discussed below question whether nudging is entirely benign, but sometimes it may be more benign than any feasible alternative.

A number of authors have found nudges problematic,[10] and there are four main concerns about nudging:

1. Nudges are disrespectful. They treat people like children.

2. Nudges tend to perpetuate and to amplify deliberative flaws.

3. Nudges undermine autonomy.

4. Nudging is condescending and arrogant.

I do not find the first two of these objections particularly powerful. If most people are myopic or subject to framing effects, then policy makers are not singling anyone out when they structure choices so that people's myopia or their susceptibility to framing either assists them to choose well or at least does not hinder them from doing so. Does respect require that policy makers pretend that people do not have deliberative crochets? If certain flaws are characteristic of human beings, is there anything disrespectful in recognizing and making use of them? Just as we put handrails and abrasive strips on steps because

7 C. Sunstein, op. cit., p. 446.
8 Ibid., p. 416.
9 Ibid., p. 431.

10 See, for example, S. Conly, op. cit.; T. Grüne-Yanoff, R. Hertwig, op. cit.; D. Hausman, B. Welch: To Nudge or Not to Nudge, in: Journal of Political Philosophy, Vol. 18, No. 1, 2010, pp. 123-36; and J. Waldron: It's All for Your Own Good, in: New York Review of Books, Vol. 61, No. 15, 9 October 2014, available at http://www.nybooks.com/articles/2014/10/09/cass-sunstein-its-all-your-own-good/.

we recognize that people are prone to stumbling, so we assign defaults to prevent the deliberative analogue to a stumble. The case would be different if government policy singled out some social group (other than children and others with limited competence) as particularly prone to bad choices.

The second objection, that nudges keep people from learning how to avoid deliberative mistakes and that they encourage bad deliberative habits, rests on an empirical claim for which I have not seen much evidence. I am inclined to agree with Jeremy Waldron when he 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."[11] But even if informing and training people, if workable, were better than nudging them, nudging may be a good thing.

The third objection – that nudging undermines autonomy – points to a serious worry. By "autonomy", I mean the control individuals have over their own evaluation, deliberation and choice. Compare the following two policies. Employer 1 sets up a voluntary retirement plan, trains employees on how to make their own evaluation of the alternatives, and then steps back and lets the employees choose. Employer 2 sets defaults and other features of the choice circumstances in order to increase future employee contributions to their retirement plans. The first employer is enabling the employees' choices without controlling them. The second is attempting to control her employees.

The reason why nudges, such as setting defaults, seem to be ethically problematic is that they aim to "push" individuals to make one choice rather than another, quite apart from engaging in rational persuasion. The employee's freedom, in the sense of what alternatives are available, is virtually unaffected, but when this "pushing" takes the form of choice architecture, their autonomy – the extent to which they have control over their own evaluations and deliberation – is diminished. Their actions derive in part from the tactics of the choice architect, rather than exclusively from their own evaluation of alternatives. When a benign employer, Marilyn, congratulates herself on engineering the situation so that the employees chose just the pension plans that she judged to be best, Marilyn is celebrating her power over her employees. At the same time (and this is the fourth criticism), nudging suggests the superiority of those who design the nudges or perhaps even their contempt for those whom they nudge. Waldron puts it this way:

> For Sunstein's idea is that we who know better should manipulate the choice architecture so that those who are less likely to perceive what is good for them can be induced to choose the options that *we* have decided are in their best interest.[12]

This threat to autonomy could be grave. But I suspect that nudges have too little power to threaten autonomy seriously. For example, a recent experiment by Volpp et al. employed a panoply of nudges to improve people's compliance with medication regimes:[13] "This was a kitchen-sink approach. It involved direct financial incentives, social support nudges, health care system resources and significant clinical management."[14] Yet it failed to improve compliance. Obviously, a single study does not show that nudging has little force. But the results are suggestive. Compared to the enormous power of alternatives, such as inciting emotions of fear and hatred, which so easily mobilize violent mobs and mass irrationality, nudges are small potatoes.

I conclude that there is little practical reason to get excited about nudges. There is little to be said in general on their behalf, except in those circumstances where the only feasible alternative methods of influencing behavior are objectionable. There is also not much to be said against them, because they pose scant threats to deliberative competence or self-respect, and it is unlikely that they will undermine autonomy in any significant degree.

On the other hand, the findings of behavioral economics that suggest a role for nudging should be alarming to normative economists, because they threaten the grounds upon which economists evaluate alternatives. If preferences, especially as shown in choices, do not reliably indicate what is good for individuals – as there was already good reason to believe before behavioral economics documented the deliberative mistakes to which people are prone – then normative economics requires a drastic rethinking.

11 J. Waldron, op. cit.

12 Ibid.

13 K. Volpp et al.: Effect of Electronic Reminders, Financial Incentives, and Social Support on Outcomes After Myocardial Infarction: The HeartStrong Randomized Clinical Trial, in: Journal of the American Medical Association: Internal Medicine, Vol. 177, No. 8, 2017, pp. 1093-1101.

14 A. Carroll: Don't Nudge Me: The Limits of Behavioral Economics in Medicine, New York Times, 6 November 2017, available at https://www.nytimes.com/2017/11/06/upshot/dont-nudge-me-the-limits-of-behavioral-economics-in-medicine.html.