

The concept of nudge and its moral significance: a reply to Ashcroft, Bovens, Dworkin, Welch and Wertheimer

Yashar Saghai

I warmly thank Richard Ashcroft, Luc Bovens, Gerald Dworkin, Brynn Welch, and Alan Wertheimer for their insightful comments on my article.¹ As I do not have the space to discuss all the questions they raise, I will focus on four concerns that run through my commentators' responses.

CLUTTERING OUR MINDS WITH TRIVIALITIES

Wertheimer argues that my Choice-Set Preservation Condition is objectionable because it wrongly implies that any reduction of the choice-set is incompatible with the preservation of freedom of choice.² He notes that removing unimportant options may enhance the influencee's 'deliberative capacities and, perhaps, her freedom'.

I agree with Wertheimer's substantive claim, although I do not think it is incompatible with my view. As Joel Feinberg argued, freedom of choice is undermined when fecund options are eliminated.³ Fecund options are options that lead to another array of options. For instance, the choice between several universities opens up different career prospects. Restricting those upstream choices has consequences on an unspecifiable number of branching downstream choices. However, restricting 'limited' options (eg, multiple toothpaste brands) that are not gateways to other options may actually facilitate decision-making and ultimately increase freedom. Cluttering our minds with trivialities would undermine the exercise of practical reason.

SUBSTANTIAL NONCONTROL AND EASY RESISTIBILITY

Welch's objections focus on my definition of easy resistibility.⁴ He claims that my definitions of easy resistibility and of the

ability to resist an influence easily are inconsistent. Dworkin argues that the language of the 'ability to easily resist' overlooks the critical distinction between having a capacity and exercising it.⁵ While I stand by my understanding of nudges, I would like to rephrase my view in order to consolidate some of the definitions I offer, taking into account the comments of Dworkin and Welch.

I defend the thesis that an influence preserves freedom of choice if and only if it is choice-preserving and is either fully or substantially noncontrolling. In my article, I focus on substantial noncontrol because nudges are best conceived as influences that work primarily by activating shallow cognitive processes, yet are choice-set preserving and substantially noncontrolling.

The main challenge I face is to offer a compelling analysis of the Substantial Noncontrol Condition. I suggest understanding this concept in the following modified terms:

A's influence to get B to ϕ is substantially noncontrolling if B can easily resist A's attempt to get her to ϕ .

I maintain that B can easily resist A's influence if and only if B is not subject to influences, or put in circumstances that would either significantly undermine the

Correspondence to Yashar Saghai, MA, PhD, Berman Institute of Bioethics, Johns Hopkins University, 1809 Ashland Avenue, Baltimore, MD 21205, USA; ysaghai@jhu.edu

relatively effortless exercise of the capacities that enable B to oppose A's pressure to get her to ϕ , or weaken those very capacities.

The empirical literature in cognitive science identifies at least two sets of cognitive capacities necessary for such form of control over one's behaviour, choices, or decisions: the capacities for goal-conflict recognition and the capacities for goal-conflict resolution.⁶ This terminology puts what I label in my article 'attention-bringing' and 'inhibitory' capacities in sharper form. The aim of models of control that distinguish between these two sets of capacities is to describe how and when lower-level cognitive processes are overridden by higher-level cognitive processes, enabling individuals to control their behaviour and to allocate their finite cognitive resources optimally.

Goal-conflict recognition always involves processes that monitor the output of various lower-level cognitive processes to detect a conflict or a mismatch between the agent's explicit or implicit goal and her behavioural response. Examples include situations in which our non-deliberative cognitive processes fail to lead us smoothly to an adaptive response to our environment. This is the case when we ride a bicycle on autopilot and suddenly notice an unusual problem on the road that cannot be addressed if we rely on our automatic skills. Similarly, sometimes reliance on mental heuristics to solve a problem does not result in a sufficiently good answer to the problem we try to address. Optimally, processes involved in goal-conflict recognition send a signal to initiate higher-level processes, such as more formal and systematic reasoning. Goal-conflict recognition processes can be either conscious⁷ or nonconscious.^{8,9}

In my article, I stress the role of conscious processes. I argue that at least when there is sufficient conflict between the influencer's and the influencee's goals, the influencee's goal-conflict recognition processes might be activated, and the influencee may be in a position to deliberate over her choice and resist the attempt to modify the course of her behaviour. My aim in that piece is to convince those who maintain that control over our choices can only be the result of conscious deliberation that they have no reason to believe that influences activating shallow cognitive processes are always incompatible with freedom of choice, as they understand it. But of course, we do not have to agree with their claim that control needs to be

conscious. This is the point I argue next, and which I do not explicitly address in my article.

According to some cognitive scientists, certain goal-conflict recognition processes can be entirely nonconscious.⁷ These processes are flexible and efficient in performing various functions, including goal-pursuit in the face of distractions or obstacles in the environment. Experiments in non-human animals have, for instance, shown that mammals (for instance, rats) can delay gratification by pursuing a distant goal and inhibiting their prepotent responses to a stimulus, without relying on conscious deliberation. Similarly, humans often achieve self-control through the use of automatised habits, rather than through deliberation.⁹

Naturally, the capacity to resist an influence easily does not solely require conscious or nonconscious goal-conflict recognition but also goal-conflict resolution. I refer here in particular to our capacity to inhibit our propensity to do what the influencer wants us to do. Sometimes we are in a position to resist an unwanted influence easily (by recognising a conflict and inhibiting a response), but also need to find a better solution to the problem at hand. For instance, suppose that a political campaign aims at discrediting a candidate by spinning statistics. You notice that there is something not quite right there and you are able easily to inhibit your propensity to believe what you are told. However, you are still left with a problem if you do not know enough about statistics to come up with a better way to process the statistical data. You may fall back on using a misleading heuristic and form a false belief or endorse a true belief without having any justification for it.

If this view holds, then nudges that either facilitate or counter the action the influencee wants to perform are often easily resistible. Nudges that shape the influencee's actions *de novo* may or may not be easily resistible (Welch is right on this point).

However, *contra* Welch, I do not see why on my account libertarian paternalism is a not a candidate justificatory strategy for the defense of certain nudges. However, I claim that libertarian paternalism justifies, at best, a subclass of nudges, namely those that are truly paternalistic (or are at least performed solely for the benefit the nudgees), and whose aim is endorsed by the influencee or would be endorsed by her if she were fully rational. If few nudges satisfy these conditions, we should conclude that libertarian

paternalism is not as compelling a justificatory strategy for the defense of nudges as it claims to be, and needs to be supplemented by alternative strategies.

NUDGES AND DECEPTION

I maintain that a substantially noncontrolling influence should not create circumstances that weaken the capacities necessary to resist an influence easily or undermine the exercise of those capacities. Dworkin is right when he notes that deception does not weaken our capacities to resist an influence easily, but he also concedes that the deception may undermine the exercise of those capacities.⁵ I agree with Dworkin that whether a particular instance of deception is easily resistible is an empirical question, not a conceptual one.

Bovens takes up this empirical question when he writes that in *Asparagus-Lovers*, 'I can easily recognise that pressure is being exercised when I am told that I used to like asparagus', although I may not be cognizant of the falsity of the information about my taste.¹⁰ I disagree. The experimental design of *Asparagus-Lovers* is intended to mislead subjects into believing that no one is trying to influence their food preferences. The questionnaires subjects had to fill out were allegedly part of a study of the relationship between food preferences and personality.¹¹

Things are more complicated in the version of *Less Than You Think* that is based on deceptive information. As Kyle Whyte and Evan Selinger note, 'people perceive and interpret meaning in situations where they have to make choices' (p. 930).¹² The bar for whether an influence is likely to undermine the exercise of the average person's capacities to resist an influence easily may be set higher when the average person can reasonably be expected to anticipate being influenced in a particular direction and has the required capacities to resist that influence. I wholeheartedly concur with Ashcroft's view that future research on the ethics of nudging needs to explore the problem of trust and authority in order to determine when nudges are legitimate.¹³

THE MORAL SIGNIFICANCE OF THE NUDGE/PROD DISTINCTION

Wertheimer and Bovens wonder how much moral work the nudge/prod distinction really does.^{2,10} Here, it is important to avoid a potential misinterpretation of my thesis. I do not claim that substantial non-control is a right-making feature and substantial control a wrong-making feature of an influence. I argue that the degree of control an influence exerts on us is a

morally relevant characteristic of that influence. My aim is to elaborate an account of nudges (and of influences more generally) that can be used by proponents of conflicting theories of the moral significance of control over one's choices.

Those who believe that control over one's choices always matters will consider that prods should be avoided, but that nudges are always preferable to prods. The degree of control an influence exerts on us is obviously an important question for those who care equally about the protection of all liberty interests.

My own view is that all liberties are not on a moral par. I maintain elsewhere that public health authorities are entitled to use certain substantially or fully controlling means to interfere with some choices in order to promote public welfare without added justificatory burden.¹⁴ In my account, the degree of control matters morally when an influencer interferes with liberties that play a vital, or at least a consequential role in enabling us to lead lives that are, in their broad contours, self-determining. I do not claim that interference with these important liberties is always impermissible, however. My view does not entail that interfering with reproductive choices or parental freedom cannot be justified at all, or that policies compromising present self-determination

for the sake of future self-determination should be ruled out.^{2–10} Rather, I claim that those who wish to interfere with liberties that deserve to be strongly protected have to shoulder a heavier burden of justification to make the case for the policies they advocate. There is no such added burden of justification when public health authorities interfere with non-consequential liberties.¹⁴ Nudges are therefore not systematically preferable to more controlling influences.

Competing interests None.

Provenance and peer review Commissioned; internally peer reviewed.

To cite Saghai Y. *J Med Ethics* 2013;**39**:499–501.

Received 4 December 2012

Accepted 17 December 2012

Published Online First 20 February 2013



- ▶ <http://dx.doi.org/10.1136/medethics-2012-100727>
- ▶ <http://dx.doi.org/10.1136/medethics-2012-101060>
- ▶ <http://dx.doi.org/10.1136/medethics-2012-101061>
- ▶ <http://dx.doi.org/10.1136/medethics-2012-101109>
- ▶ <http://dx.doi.org/10.1136/medethics-2012-101110>
- ▶ <http://dx.doi.org/10.1136/medethics-2012-101111>

J Med Ethics 2013;**39**:499–501.

doi:10.1136/medethics-2012-101112

REFERENCES

- 1 Saghai Y. Salvaging the concept of nudge. *J Med Ethics* 2013;**39**:487–93.
- 2 Wertheimer A. Should 'nudge' be salvaged? *J Med Ethics* 2013;**39**:498–9.
- 3 Feinberg J. *Harm to others. The moral limits of the criminal law*, Vol. 1. New York: Oxford University Press, 1984: 208.
- 4 Welch B. Shifting the concept of nudge. *J Med Ethics* 2013;**39**:497–8.
- 5 Dworkin G. Lying and nudging. *J Med Ethics* 2013;**39**:496–7.
- 6 Stanovich KE. *Rationality and the reflective mind*. New York: Oxford University Press, 2011:79.
- 7 Thompson VA. Dual-process theories: a metacognitive perspective. In: Evans JStBT, Frankish K. *In two minds: dual processes and beyond*. New York: Oxford University Press, 2009: 171–96.
- 8 Suhler CL, Churchland PS. Control: conscious and otherwise. *Trends Cogn Sci* 2009;**13**:341–7.
- 9 Faucher L. Tirer la responsabilité au clair: le cas des attitudes implicites et le révisionnisme. *Les Ateliers de l'Éthique/The Ethics Forum* 2012;**7**:179–212.
- 10 Bovens L. Why couldn't I be nudged to dislike a Big Mac? *J Med Ethics* 2013;**39**:495–6.
- 11 Laney C, Morris EK, Bernstein DM, et al. Asparagus, a love story: healthier eating could be just a false memory away. *Exp Psychol* 2008;**55**:291–300.
- 12 Selinger E, Whyte K. Is there a right way to nudge? The practice and ethics of choice architecture. *Sociol Compass* 2011;**15**:923–35.
- 13 Ashcroft R. Doing good by stealth: comments on 'Salvaging the concept of nudge'. *J Med Ethics* 2013;**39**:494.
- 14 Powers M, Faden R, Saghai Y. Liberty, Mill, and the framework of public health ethics. *Public Health Ethics* 2012;**5**:6–15.