

‘Better off, as judged by themselves’: a reply to Cass Sunstein

Robert Sugden¹ 

Received: 20 June 2017 / Accepted: 25 June 2017
© The Author(s) 2017. This article is an open access publication

Abstract This paper is a reply to Sunstein’s comment on my paper ‘Do people really want to be nudged towards healthy lifestyles?’ The central claim of that paper was that, in their book *Nudge*, Thaler and Sunstein switch between two different interpretations of the ‘better off, as judged by themselves’ criterion, and that consistent use of one or other interpretation would have blunted the persuasive power of the book. In this reply, I defend that claim against Sunstein’s counter-arguments.

Keywords Sunstein · Nudges · Health · Paternalism · Self-control

JEL Classification D6 · I1

The paper on which Cass Sunstein comments is an expanded version of a keynote conference lecture (Sugden 2016). My intention as a speaker was to make a serious methodological point about the justification of nudges, but in an informal and provocative way. I thank Sunstein for responding constructively to the provocation, and for doing so in good humour—recognising the significance of the question the paper addresses while excusing its lack of empirical rigour.

That question is about the principles by which, according to their advocates, nudges can be justified. Specifically, it is about the meaning of the claim, repeatedly made by Thaler and Sunstein (2008) in *Nudge*, that their recommendations are designed to ‘make choosers better off, *as judged by themselves*’ (p. 5, italics in original). I argued that Thaler and Sunstein switch between two different interpretations of the ‘as judged by themselves’ (AJBT) clause and that consistent

✉ Robert Sugden
r.sugden@uea.ac.uk

¹ School of Economics, University of East Anglia, Norwich NR4 7TJ, UK

use of one or other interpretation would have blunted the persuasive power of their book. In his comment, Sunstein tries to clarify the intended meaning of AJBT.

According to Sunstein: ‘Choice architects might well have their own ideas about what would make choosers better off, but in our view, the lodestar is people’s own judgments. To be a bit more specific: The lodestar is welfare, and people’s own judgments are an excellent way to test the question whether nudges are increasing their welfare’ (Sunstein 2017). The first sentence gives the most natural reading of AJBT—that, when deciding whether and how to nudge, the choice architect’s criterion should be to respect each chooser’s judgements about his own welfare, whether or not she (the choice architect) shares those judgements. However, the second sentence might be read as saying that the criterion is the chooser’s *actual* welfare and that this should be assessed by the most competent judge; as a matter of contingent fact, that is usually the chooser himself but, in principle, it might be the choice architect, informed by expert advice. That reading would open Thaler and Sunstein to the charge that Thaler (2015, pp. 325–326) so emphatically denies, of ‘thinking that we know what is best for everyone’. I suggest that the best way to resolve this tension, given the underlying logic of *Nudge*, is to treat the first sentence as stating the *criterion* that choice architects should be required to use, and the second sentence as providing a *reason* for imposing that requirement.

In explaining the implications of AJBT, Sunstein distinguishes between three categories of cases.

In cases in category (1), choosers have ‘clear antecedent preferences’; nudges merely provide information that choosers need in order to implement those preferences (as in the case of Elizabeth), or remind them of intentions about which they may have forgotten (as in the case of Luke). I agree that in the cases that Sunstein has in mind here, AJBT is unproblematic. However, I am not sure that *antecedent* preference is the most useful concept to use when thinking about nudges. As normally understood in economics, a preference is a general disposition to choose one thing rather than another; it applies across a whole range of actual and counterfactual decision problems. (In the canonical neoclassical model, it applies across all possible decision problems.) The crucial property of ‘behavioural’ preferences—the property that makes nudging possible—is not that preferences change over time; it is that they are *context-dependent*. A nudge does not change a person’s preference from ‘antecedent’ to ‘posterior’; it changes the context in which *given* (but possibly context-dependent) preferences are activated. I would prefer to say that in category (1) cases, choosers have *context-independent* preferences, and that the choice architect’s criterion is the satisfaction of those preferences.

In cases in category (3), choosers’ preferences are context-dependent. In Sunstein’s example, Thomas’s preference between having and not having an operation varies according to whether his attention is directed towards the potential benefits of the operation or towards its potential risks. A choice architect can affect Thomas’s choice by choosing how to present given information about benefits and risks. The problem for the AJBT criterion is that Thomas’s judgement about what makes him better off is itself context-dependent, and so cannot be used to determine the context in which he should choose.

In response to the question of what the choice architect ought to do in such cases, Sunstein concludes that 'there is no escaping at least some kind of welfarist analysis'—that is, an analysis that makes 'direct inquiries into people's welfare' (Sunstein 2017). In his comment, Sunstein does not say much about how a person's welfare is defined or assessed, but many of the arguments in *Nudge* imply that the method of enquiry is to try to reconstruct the (assumedly context-independent) *latent* preferences that fully informed choosers would reveal in the absence of psychologically induced error. Sunstein seems to endorse this methodology when he says: 'It is psychologically fine to think that choosers have antecedent preferences, but that because of a lack of information or a behavioural bias, their choices will not satisfy them' (Sunstein 2017). Here I disagree. In the first part of my paper, which summarises a fuller analysis presented by Infante et al. (2016), I argued that it is *not* psychologically fine to assume that human choices result from interactions between context-independent latent preferences and behavioural biases. I maintain that the concept of latent preference is psychologically ungrounded.

I have interpreted AJBT, as applied to category (3) cases, as referring to the judgements implicit in choosers' latent preferences. In his comment, Sunstein offers a different interpretation—that the relevant judgements are implicit in choosers' actual posterior preferences. Take the case of Thomas and the operation. We are told that, in whichever direction Thomas is nudged, he will be 'happy' with his decision, judging himself to be better off than if he had chosen differently (Sunstein 2017). In other words, any nudge that causes Thomas to change his decision can be said to make him better off, *as judged by himself*. I think Sunstein is going astray here by thinking of nudges as causing changes in preference. Suppose the doctor directs Thomas's attention towards the benefits of the operation and advises him to have it. Thomas accepts this advice. *At the moment of choice*, Thomas is thinking about the options in the frame provided by the doctor, and so he thinks he is making the right decision. But suppose that, shortly before he is wheeled into the operating theatre, he looks at some medical website that uses the opposite frame. If his preferences are context-dependent, he may now wish he had chosen differently. Sunstein is not entitled to assume that, after choosers have been nudged, their judgements become context-independent. If the AJBT criterion is to have bite—if, as Sunstein says, it is to '*discipline the content of paternalistic interventions*' (Sunstein 2017)—it must adjudicate between the judgements that the chooser makes in different contexts. That is why Thaler and Sunstein need the concept of latent preference—with all its problems.

Sunstein's category (2) contains the cases that were the focus of my paper. These are cases in which a chooser (such as Mary, the student who drinks a lot) has a self-acknowledged self-control problem: she has a continuing preference to choose in a certain way, but sometimes feels unable to resist temptations to do otherwise. In my paper, I argued that these cases allow a coherent interpretation of AJBT, but questioned how common they are, pointing to the many other psychological mechanisms that can lead people to choose contrary to expert advice about what is good for them.

I now see that the title of my paper, 'Do people really want to be nudged towards healthy lifestyles?' was ambiguous. It might be read as asking about political

opinions: How far do people approve of nudges to which they, along with their fellow-citizens, will be subject? Sunstein is right to say that, in both Europe and the US, there is widespread political support for nudges that promote choosers' own health. However, I think he overstates the evidence for the claim that people 'would like to be nudged', as opposed to believing that nudges are good for other people and accepting being nudged themselves as a small price that has to be paid for this. The paper he cites in support of this claim finds only that approval of nudges is unaffected by whether nudges that will in fact affect everyone are described in terms of their effects on 'people' or on 'you' (Jung and Mellers 2016). Conflicting findings are reported by Cornwell and Krantz (2014).

I take it that Sunstein and I are contributing to a public debate about the merits of nudging. We are not counting existing opinions; we are presenting citizens with reasons for approving or disapproving of nudges. My paper was concerned with one putative reason for approving of them—the claim that individuals want to make the choices they will be nudged towards, but believe that they fail to make those choices because of weakness of will. I accept that Sunstein's Mechanical Turk survey investigates this question. But by asking respondents whether, scanning over all domains of life, they have *any* issue of self-control, *large or small*, it sets the bar very low. And the opening sentence of Sunstein's survey question, 'Many people believe...' is a classic nudge in itself, similar to the one that is supposed to lead Mary to drink less. I think a better test of the prevalence of self-acknowledged self-control problems would be to investigate the take-up of options that are explicitly presented as self-constraint mechanisms, such as 'panic buttons' that allow online gamblers to lock themselves out of their accounts for fixed periods. (According to a recent news item about a British betting firm with 2.7 million online users, the availability of a panic button led to the locking of about 3000 accounts per week.¹) But Sunstein is right to say that I have not explored the data.

In my paper, I suggested that Thaler and Sunstein's (2008) emphasis on the AJBT criterion 'reveals a sensitivity to the criticism that their position is unacceptably paternalistic' (Sugden 2016). Sunstein says that the criterion was not 'designed to counter the charge of paternalism' and that he and Thaler 'embrace paternalism' (Sunstein 2017). I accept that Thaler and Sunstein are open about the paternalism of their approach—after all, they call it 'libertarian paternalism'. But I still think that sensitivity to potential criticism from anti-paternalists is a prominent feature of *Nudge* and that appeals to AJBT are among the ways by which Thaler and Sunstein try to head off this criticism. How else should one interpret the sentence in which 'as judged by their own preferences' is immediately followed by 'not those of some bureaucrat' (2008, pp.9–10)? A similar sensitivity is revealed in Thaler and Sunstein's insistence that unobjectionable forms of paternalism can be found everywhere. Sunstein's claim that GPS devices are paternalistic (Sunstein 2017) is an example of this line of argument. If all voluntarily hired expertise (the GPS

¹ This item appeared in *The Independent* on 23 March 2016 and at the web address <http://www.independent.co.uk/news/uk/home-news/william-hill-betting-firm-complains-panic-button-allowing-gamblers-to-take-break-is-hurting-profits-a6948766.html>.

device, the road atlas, the car mechanic, the chef, the gardener) is paternalistic, we can all embrace paternalism, but the concept loses any useful meaning.

Acknowledgements I thank Gerardo Infante for comments on an earlier version of this paper. My work is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme, Grant agreement No. 670103.

Compliance with ethical standards

Conflict of interest The author declares that he has no conflicts of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Cornwell J, Krantz D (2014) Public policy for thee, but not for me: varying the grammatical person of public policy justifications influences their support. *Judgm Decis Mak* 5:433–444
- Infante G, Lecoutey G, Sugden Robert (2016) Preference purification and the inner rational agent: a critique of the conventional wisdom of behavioural welfare economics. *J Econ Methodol* 23:1–25
- Jung J, Mellers B (2016) American attitudes toward nudges. *Judgm Decis Mak* 11:62–74
- Sugden R (2016) Do people really want to be nudged towards healthy lifestyles? *Int Rev Econ* 64(2):113–123. doi:[10.1007/s12232-016-0264-1](https://doi.org/10.1007/s12232-016-0264-1)
- Sunstein C (2017) 'Better off, as judged by themselves': a comment on evaluating nudges. *Int Rev Econ*. doi:[10.1007/s12232-017-0280-9](https://doi.org/10.1007/s12232-017-0280-9)
- Thaler R (2015) Misbehaving: how economics became behavioural. Allen Lane, London
- Thaler R, Sunstein C (2008) Nudge: improving decisions about health, wealth, and happiness. Yale University Press, New Haven