

Do people really want to be nudged towards healthy lifestyles?

Robert Sugden¹ 

Received: 10 June 2016 / Accepted: 10 September 2016 / Published online: 20 September 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract This paper assesses Thaler and Sunstein’s claim that policies that nudge individuals towards healthy lifestyles promote the welfare of those individuals, *as judged by themselves*. I argue that Thaler and Sunstein switch between two different interpretations of that clause. One interpretation gives the clause a wide range of applicability, but drains it of its content as a repudiation of paternalism. The other interpretation makes it more meaningful to say that people want to make the choices they are being nudged towards, but applies to a much narrower range of cases than Thaler and Sunstein have in mind.

Keywords Nudges · Health · Paternalism · Self-control

JEL Classification D6 (welfare economics) · I1 (health)

People can expect longer and healthier lives if they make certain kinds of lifestyle choices. This is known not only to health specialists, but also to the general public. The considered advice of medical and public health experts is readily available in forms that are designed to be easily understood. In the UK, for example, you can consult an accessible website maintained by the National Health Service, with the name ‘NHS Choices’ and the banner ‘Your health, your choices’. (The emphasis on individual choice is surely intentional, despite the rather prescriptive tone of the contents.) A few minutes browsing will be enough for you to learn that you are advised to have a diet that is low in saturated fats, salt, sugar and processed meats and high in wholegrain carbohydrates, pulses, fruit, vegetables and oily fish, that

✉ Robert Sugden
r.sugden@uea.ac.uk

¹ School of Economics, University of East Anglia, Norwich NR8 6SL, UK

you should drink no more than small amounts of alcohol, that you should take regular exercise, and that you should not smoke. Simple guidelines are provided: you should eat at least five portions of a variety of fruit and vegetables a day; you should not drink more than 14 units of alcohol over the course of a week; you should try to have at least 150 min of physical activity each week. These guidelines are widely known and are frequent topics of media attention. Nevertheless, many people do not follow them.

Behavioural economists often argue that governments should respond by *nudging* people towards healthy lifestyles. The idea is to design choice architecture (that is, the environments in which choices are made) in ways that engage with psychological mechanisms that economists have traditionally regarded as non-rational, but which are known to influence actual choices. These mechanisms are to be used to increase the likelihood of healthy lifestyle choices while not restricting other opportunities. The current paper is about one common justification for such nudges—the claim that *individuals want to make the choices that they are being nudged towards*.

I am interested in this claim because I am trying to develop a general form of normative economics that is compatible with behavioural findings. As part of this enterprise, I want to understand and appraise the principles that behavioural economists use when they do normative analysis. This is not easy, because behavioural economists are often rather casual in normative argument. One fixed point is that when nudge policies are being justified, the claim that people want to be nudged is a recurring theme. I want to know whether this claim is coherent, whether it is consistent with the available evidence, and whether it is consistent with the policy recommendations that it is supposed to justify.

Let me emphasise that this paper is concerned only with this one specific argument in favour of nudges. The nudge literature contains many other arguments. For example, there is the claim that individuals' welfare would be greater with healthier lifestyles, whatever they may believe, and that this is a sufficient reason for nudging. There is the argument that public policy favours the adoption of healthy lifestyles and that the job of the economist is merely to find out 'what works' in achieving given objectives. And there is the argument that one person's poor health imposes external costs on others and that healthy lifestyle nudges are justified in terms of the interests of those who bear those costs. I will not be concerned with such arguments. Nor will I be concerned with arguments *against* nudging, such as the claim that nudges are ineffective in the long run, or the claim that they compromise individual autonomy. All these arguments for and against nudging are important and call for scrutiny. I have expressed my views about them in other papers (Sugden 2008, 2009; Infante et al. 2016). But it is easier to think clearly if one takes one argument at a time. I will focus on how the argument that people want to be nudged is deployed by Thaler and Sunstein in their book *Nudge* (2008).

Thaler and Sunstein go out of their way to claim that the criterion by which they assess nudges is that of making individuals better off, not as judged by the policy maker, social planner or 'choice architect' who implements the nudge, but as judged by the person who is being nudged. This is stated explicitly in their opening chapter: their recommendations for nudging, they say, are designed to 'make choosers better

off, *as judged by themselves*' (p. 5; italics in original). A few pages on, after claiming that the assumption that people usually choose in their best interests is 'obviously false', they say: 'So long as people are not choosing perfectly, some changes in the choice architecture could make their lives go better (as judged by their own preferences, not those of some bureaucrat)' (pp. 9–10). Later in the book, using one of their favourite distinctions between 'Econs' (individuals as traditionally modelled in neoclassical economics) and 'Humans', they say:

Some firms sell cigarettes; others sell products that help you quit smoking. Some firms sell fast food; others sell diet advice. If all consumers are Econs, there is no reason to worry about which of these competing interests wins. But if some of the consumers are Humans who sometimes make bad choices (as judged by themselves, of course), then all of us may have an interest in which set of firms wins the battle (p. 80).

The importance of the 'as judged by themselves' clause is underlined rather tetchily by Thaler in his recent book *Misbehaving* (2015). Responding to criticisms of *Nudge*, Thaler emphasises 'a point that critics of our book seem incapable of getting'. That point, we are told, is that Thaler and Sunstein 'have no interest in telling people what to do. We want to help them achieve their *own* goals'. As proof that *Nudge* does not tell people what to do, Thaler points to the first '*as judged by themselves*' clause in that book:

The italics are in the original but perhaps we should also have used bold and a large font, given the number of times we have been accused of thinking that we know what is best for everyone. ... We just want to reduce what people would themselves call errors (pp. 325–326).

It is easy to see the role of this clause in Thaler and Sunstein's argumentative strategy. Their eagerness to disassociate themselves from 'bureaucrats' who 'tell people what to do' reveals a sensitivity to the criticism that their position is unacceptably paternalistic [rather than, as they maintain, paternalistic in such a 'weak, soft and nonintrusive' sense that even libertarians cannot object to it (Thaler and Sunstein 2008, pp. 4–6)]. Paternalism can be defined in different ways, but from now on I will use the word in the sense defined by Le Grand and New (2015, p. 23): 'A government intervention is paternalistic towards an individual if it is intended to address a failure of judgement by that individual [and to] further the individual's own good'. This definition makes paternalism a matter of *intention*: its key feature is the substitution of the government's judgement for that of the individual on the grounds that the individual's judgement of her own good is in error. It allows us to see why some readers find Thaler and Sunstein's programme objectionably paternalistic, even if it does not impose any substantial restrictions on individual choice.

But what exactly does the clause mean, and does it really make Thaler and Sunstein's position non-paternalistic? I will argue that Thaler and Sunstein switch between two different interpretations of this clause. One interpretation gives it a wide range of applicability, but drains it of its content as a repudiation of paternalism. The other interpretation gives more meaning to the claim that people

want to make the choices they are being nudged towards, but applies to a much narrower range of cases than Thaler and Sunstein have in mind. The rhetorical success of *Nudge* depends heavily on this ambiguity.

1 Behavioural welfare economics

In the passages I have quoted, Thaler and Sunstein are declaring that their normative criterion is based on individuals' *judgements* about what makes them better off, and about what their goals are. There is a clear implication that these judgements are *subjective*: different individuals in similar circumstances may make different judgements about what is good for them, without those differences showing that anyone's judgement is in error. The policy maker is expected to respect those judgements, rather than to implement her own judgements about what is good for people, or to implement the judgements of some consensus of experts. But how are individuals' subjective judgements defined, and how is the policy maker (or the normative economist who advises her) to elicit them?

Immediately after presenting the principle of trying to make choosers 'better off, as judged by themselves', Sunstein and Thaler promise:

[We will] show that in many cases, individuals make pretty bad decisions—decisions that they would not have made if they had paid full attention and possessed complete information, unlimited cognitive abilities, and complete self-control (2008, p. 5).

The obvious interpretation is that individuals' judgements about their own welfare are revealed in the decisions that those individuals would have made in the absence of error—that is, with full attention, complete information, unlimited cognitive ability and complete self-control. The normative economist's job is to reconstruct individuals' error-free *latent preferences* and to recommend policies that satisfy those preferences.

This approach of *behavioural welfare economics* is widely used by prominent behavioural economists (Bleichrodt et al. 2001; Camerer et al. 2003; Köszegi and Rabin 2007; Bershears et al. 2008; Salant and Rubinstein 2008; Bernheim and Rangel 2009) and has been endorsed from the perspectives of economic philosophy (Hausman 2012) and the economics of social policy (Le Grand and New 2015). In a recent paper, Gerardo Infante, Guilhem Lecouteux and I have documented the extent to which this approach is used, reconstructed its logic, and subjected it to a methodological critique (Infante et al. 2016). We argue that its implicit model of a human decision maker is that of an *inner rational agent* interacting with the world through an error-prone *psychological shell*. The inner agent is imagined to have stable, context-independent preferences of a broadly neoclassical kind. These are the latent preferences that policy makers should try to satisfy, and which express the individual's judgements about what is good for her. Psychological factors appear in the model only as obstacles that can prevent the inner agent from achieving its goals—obstacles such as inattention, cognitive overload and lack of self-control.

Despite Thaler and Sunstein's label, the decision maker described by this model is not a Human in the ordinary sense of the word. It is a faulty Econ.

Infante, Lecouteux and I argue that this model lacks adequate psychological foundations. The existence and rationality of latent preferences is assumed, not explained. Latent preferences are defined as the subjective judgements that a person's choices would reveal in the absence of error, but we are given no independent criterion by which errors can be identified. Nor are we given any reason to suppose that, after the effects of specific errors have been filtered out, individuals' choices will be free of the 'anomalies' that are the *raison d'être* of behavioural welfare economics. Take the familiar case of Thaler and Sunstein's cafeteria, in which customers' choices between food items vary according to the display chosen by the cafeteria director. Thaler and Sunstein's recommendation is that the director displays the items so that customers tend to choose as they would have done, had their reasoning not been distorted by error. But if error-free choices are affected by the display too, we are no further forward.

Because of this gap in the analysis, the claim that behavioural welfare economics uses individuals' own judgements about their welfare often turns out to have little content. For example, here is how, as part of their discussion of the differences between Humans and Econs, Thaler and Sunstein justify nudges towards healthy eating:

Consider the issue of obesity. Rates of obesity in the United States are now approaching 20 per cent, and more than 60 per cent of Americans are considered either obese or overweight. There is overwhelming evidence that obesity increases risks of heart disease and diabetes, frequently leading to premature death. It would be quite fantastic to suggest that everyone is choosing the right diet, or a diet that is preferable to what might be produced with a few nudges. ... We do not claim that everyone who is overweight is necessarily failing to act rationally, but we do reject the claim that all or almost all Americans are choosing their diet optimally... With respect to diet, smoking, and drinking, people's current choices cannot reasonably be claimed to be the best means of promoting their well-being. Indeed, many smokers, drinkers and overeaters are willing to pay third parties to help them make better decisions (p. 7).

Notice that Thaler and Sunstein are not identifying specific errors of reasoning that are being made by obese Americans. The argument goes in the other direction. Because it is (supposedly) so obvious that obesity is contrary to a person's best interests, we are invited to infer that obese Americans must be reasoning erroneously—that their ability to act on their latent preferences must somehow be being obstructed. Apart from the final sentence (which surely refers only to a small minority of the obese population), there is no argument that healthy diets would be good for the obese, *as judged by themselves*, or that the obese are aware of the errors they are supposedly making. The argument is straightforwardly paternalistic. That is not to say that it would be wrong for policy makers to try to nudge people towards healthier eating: this paper is not about whether paternalism is justified. My point is that the 'as judged by themselves' clause is not achieving what I take to be its

purpose, of countering the criticism that nudge policies are paternalistic. Or at least, it does not achieve this purpose if that clause is interpreted as it is in behavioural welfare economics.

2 Akrasia

The remark about people being willing to pay third parties to help them make better decisions may suggest a different reading of ‘as judged by themselves’. On this alternative reading, latent preferences are not a hypothetical construct, defined by what the individual would prefer if ideally rational. Instead, they are the preferences that the individual *actually endorses* in some independently definable circumstances (perhaps in a cool emotional state, reflecting about her welfare), and which in some sense she continues to acknowledge even when she fails to act on them. In other words, Thaler and Sunstein may be using a model of self-acknowledged *akrasia* (that is, failure of self-control).

Some of the cases they discuss certainly fit this interpretation. For example, their chapter on ‘Resisting Temptation’ begins with a story about a dinner party hosted by Thaler. He has put out a large bowl of cashew nuts for pre-dinner nibbles. Realising that the nuts are being eaten very quickly and that his guests’ appetites for dinner will be impaired, he removes the bowl to the kitchen, where it is out of sight. Crucially for the story, the guests immediately thank him for doing this (pp. 40–41). If we take it that the guests are free to walk into the kitchen, Thaler’s action is a paradigm case of a nudge (and is very similar to the case of the cafeteria director who manipulates the display of food items). But the guests’ thanks add an extra dimension to the story. On the most natural interpretation, the guests were aware that the sight of the nuts induced an urge to eat they found hard to resist but which, even at the moment of eating, they felt a desire not to act on. As soon as the visual stimulus was removed, they endorsed that desire as their true preference. It is reasonable to conclude, as Thaler and Sunstein do, that the guests wanted the outcome that they were being nudged towards.

Thaler and Sunstein go on to argue that, because of what George Loewenstein (1996) has diagnosed as the ‘hot–cold empathy gap’, people often fail to anticipate the power of temptation, and so expose themselves to temptations that they mistakenly expect to be able to resist. In their story: ‘Tom is on a diet and agrees to go out to a business dinner, thinking he will be able to limit himself to one glass of wine and no dessert. But the host orders a second bottle of wine and the waiter brings by the dessert cart, and all bets are off’ (p. 42). It is easy to imagine a Tom who, even while accepting a share of the second bottle and choosing a *crème brûlée*, acknowledges that he is acting against his better judgement. Such a Tom might reasonably be said to want to be nudged towards keeping his resolutions—at least if, when accepting the dinner invitation, he was not 100 % confident of his willpower.

Thaler and Sunstein interpret *akrasia* in terms of a dual-self model, loosely associated with Daniel Kahneman’s (2003, 2011) dual-process theory of the mind (Thaler and Sunstein, pp. 19–39). In Thaler and Sunstein’s account, the two selves are the ‘Planner’ and the ‘Doer’:

You can think of the Planner as speaking for your Reflective System, or the Mr Spock lurking within you, and the Doer as heavily influenced by the Automatic System, or everyone's Homer Simpson. The Planner is trying to promote your long-term welfare but must cope with the feelings, mischief and strong will of the Doer, who is exposed to the temptations that come with arousal (p. 42).

Clearly, there is some affinity between this dual-self model of akrasia and the model of the inner rational agent that underlies behavioural welfare economics. But they should not be conflated.

Notice that many of the cases in which Thaler and Sunstein need to invoke the 'as judged by themselves' clause do not involve failures of self-control. Recall that, in reconstructing a person's latent preferences, we have to undo the effects of imperfect information and limited cognitive ability. If a person's preferences are based on the information that is available to him and on the best calculations and inferences that he is capable of making, it is hard to see how he can simultaneously think of himself as acting against his better judgement. For example, one of the central themes in Thaler and Sunstein's chapter on 'When Do We Need a Nudge?' is that 'rare, difficult choices are good candidates for nudges'. The reason for this is that one of the most effective ways in which people learn about the likely consequences of actions is through personal experience. Thus, people are liable to have false expectations about the consequences of actions that they take only rarely:

Unfortunately, some of life's most important decisions do not come with many opportunities to practice. Most students choose a college only once. Outside of Hollywood, most of us choose a spouse, well, not more than two or three times (pp. 74–75).

The same applies to actions that are taken regularly, but whose worst consequences occur only with low probability or only after a long lapse of time: 'Someone can eat a high-fat diet for years without having any warning signs until the heart attack. When feedback does not work, we may benefit from a nudge' (p. 75).

If these arguments for nudging are to be consistent with the 'as judged by themselves' clause, we need to interpret that clause in terms of *hypothetical* latent preferences, not preferences that individuals actually endorse but are sometimes tempted to act against. The point of the story of the person with the high-fat diet (let us call him Bill) is surely that he does *not* think that this diet is bad for his heart. Over many years, he has a consistent sense that he is eating sensibly. It is Thaler and Sunstein, in the role of experts, who are questioning Bill's judgement. They are appealing to the hypothetical choices of an inner rational agent whose reasoning takes account of information that the real Bill is unaware of or does not attend to, or whose significance he does not appreciate.

Thaler and Sunstein might reasonably argue that, over the years of his high-fat diet, the real Bill must have been exposed to information and experiences that provided good grounds for the belief that such a diet increases the risk of heart disease. His failure to make this inference might be interpreted *by an outside observer* as an error of reasoning. But recall Thaler's claim that he and Sunstein

‘just want to reduce what people would themselves call errors’. The outside observer is not entitled to infer that the high-fat diet was bad for the real Bill, *as judged by himself* and before he experienced his own heart attack. Thus, she is not entitled to infer that the real Bill wanted to be nudged towards eating less fat.

3 A thought experiment: the questionnaire

As an aid to thinking about the difference between the two interpretations of ‘as judged by themselves’, I offer the following thought experiment. Consider Jane, one of the customers at Sunstein and Thaler’s cafeteria. Jane is currently in good health but seriously overweight. Professional dieticians would agree that her long-term health prospects would be better if she ate less high-fat, sugar-rich food (such as the cafeteria’s cream cake) and more fruit and vegetables (such as the cafeteria’s fresh fruit option). But, despite knowing what dieticians advise, she usually chooses the cake. Imagine that, in an attempt to uncover the reasoning that leads people like Jane to ignore dietary advice, we are designing a questionnaire which asks: ‘*Which of the following statements best represents your reasons for choosing cake rather than fruit, contrary to the recommendations of health experts?*’ The thought experiment is to produce statements that Jane might plausibly assent to, and that could be used to diagnose her mode of reasoning.

There should be one statement that picks up failures of self-control:

- (a) I always go into the cafeteria having resolved not to choose cake, but when I see the cake at the front of the counter, I can’t resist the temptation.

Now for some others:

- (b) I get a lot of pleasure from eating sweet and fatty foods, and the thought of living to a great age doesn’t appeal to me.

The first part of this statement describes a plausible taste, but the second part may pick up an empathy gap between age groups, compounded by lack of experience of the sources of happiness open to old people.

- (c) When I am a few years older I will adopt a healthier diet, so my current eating habits are not a problem.

This statement may pick up procrastination (planning to change one’s behaviour in a future that is constantly being deferred) or an empathy gap (the respondent mistakenly imagines that her older self will be less susceptible to temptation, or will get less pleasure from food, than her present self does).

- (d) The expert advice sets unrealistic standards. Most of the people I know eat at least as much sugar and fat as I do.

This statement picks up the heuristic of *social proof*, by which people match their behaviour to that of others, even if they cannot rationalise what those others are doing (Cialdini 1984). It may also pick up a self-serving bias in the selection of comparators.

(e) All my grandparents were thin but died relatively young. It's quite likely that I will die young too, whatever I eat.

(f) All my grandparents were fat but lived to ripe old ages. It's quite likely that I will have a long life too, whatever I eat.

These statements pick up the heuristic of *availability*, which induces a tendency to overweight personal experience relative to statistical evidence and to reports, however reliable, of other people's experiences (Tversky and Kahneman 1973).

(g) Expert health advice is always changing; in a few years, experts may be recommending high-fat, high-sugar diets.

Expert health advice *does* change over time, but scepticism about the robustness of unwelcome recommendations may be a way of reducing *cognitive dissonance* by adjusting one's beliefs to fit one's preferences (Festinger 1957).

(h) Whatever I eat, I put on weight, so for me there is no point in trying to eat more healthily.

This statement may pick up self-deception about how much the respondent actually eats. It is known that people tend to under-report the calorie content of their diets (Murakami and Livingstone 2015).

The point of this thought experiment is that there are many ways in which a person can explain why she knowingly acts against well-grounded but unwelcome recommendations from experts. Failure of self-control is only one of these ways. My conjecture, based both on casual social experience and on general scientific knowledge about the prevalence of empathy gaps, procrastination, conformity, cognitive dissonance, self-deception and the tendency to overweight evidence from personal experience, is that lack of self-control would not be particularly common as a *self-ascribed* explanation; statements like (b), (c), (d), (e), (f), (g) and (h) would be more common.

Notice that each of the eight statements in my questionnaire represents a mode of reasoning, heuristic or bias that is known to be a common property of human psychology. However, (a) differs from the others in *admitting to* a failure on the part of the respondent herself: she has failed to choose as she intended (and, by implication, has failed to choose according to her own best judgement). All the other statements maintain that the respondent knows what she is doing and has chosen what, on balance, she believes is best for her. They maintain that the error is in the experts' recommendation, applied without recognition of the respondent's circumstances, not in the respondent's failure to follow it. My guess is that most people prefer not to explain their own behaviour as the result of error.

It might be objected that by including an admission of error only in statement (a), I have introduced a bias into my argument. One might reasonably claim that none of the eight statements gives a *valid* reason for not following the expert recommendation. (Indeed, one might even say that, by treating an irresistible temptation as a constraint, (a) is unique in giving a valid reason.) But suppose we rephrase (a) so that, like the other seven statements, it does not involve an admission of error, but still picks up failures of self-control:

(a') I have resolved not to eat cake very often, but if I'm feeling in the mood for cake, that's what I choose. I don't believe in being obsessive about my resolutions.

This statement allows the respondent to interpret breaking a resolution as showing spontaneity or flexibility rather than weakness of will. For that reason, respondents might be more likely to agree with (a') than with (a). An external observer might judge (a') to be a rationalisation of a failure of self-control. (Or he might not: it is entirely coherent to treat spontaneity as a desirable character trait.) But that is beside the point. My concern is with the distinction between the two interpretations of 'as judged by themselves'. What is at issue is whether people—real human beings, not their imagined inner rational agents—want to be nudged. If Jane, in the cool emotional state of a questionnaire respondent, endorses her propensity to break her own resolutions, we are not entitled to claim that her cake-eating behaviour is contrary to her own judgements about what most matters to her.

My claim is that *self-acknowledged* self-control problems are a lot less common than many behavioural economists seem to think. Even if behavioural economists or policy makers feel confident that people's lifestyle choices are based on some kind of error, they should not jump to the conclusion that the error is a self-acknowledged failure of self-control, or that (as Thaler puts it) it is what those people *would themselves call* an error.

4 Conclusion

I have argued that, in justifying policies which nudge people towards healthy lifestyle choices, Thaler and Sunstein conflate two different versions of the claim that their proposals are designed to make choosers better off, as judged by themselves. Because of this ambiguity, they can give what appear to be non-paternalistic justifications for interventions that in fact are paternalistic in intent. They can give what appears to be support for the claim that the people at whom their nudges are directed want to be nudged. But these appearances are misleading. Advocates of nudging should come clean about the paternalism of their position and defend it directly.

Acknowledgments This is an expanded version of a keynote lecture given at the HEIRS (Happiness, Economics and Interpersonal Relations) conference on 'Economics, Health and Happiness' at the University of Applied Sciences and Arts of Southern Switzerland in January 2016. I thank participants at that conference and Gerardo Infante for comments. I take this opportunity to express my appreciation of how much the HEIRS network owes to the work of the late Pierluigi Porta. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme, grant agreement No. 670103.

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

- Bernheim D, Rangel A (2009) Beyond revealed preference: choice- theoretic foundations for behavioral welfare economics. *Q J Econ* 124:51–104
- Bershears J, Choi J, Laibson D, Madrian B (2008) How are preferences revealed? *J Public Econ* 92:1787–1794
- Bleichrodt H, Pinto-Prades J-L, Wakker P (2001) Making descriptive use of prospect theory to improve the prescriptive use of expected utility. *Manage Sci* 47:1498–1514
- Camerer C, Issacharoff S, Loewenstein G, O'Donoghue T, Rabin M (2003) Regulation for conservatives: behavioral economics and the case for 'asymmetric paternalism'. *University of Pennsylvania Law Review* 151:1211–1254
- Cialdini R (1984) *Influence: the psychology of persuasion*. Quill Press, New York
- Festinger L (1957) *A theory of cognitive dissonance*. Stanford University Press, Stanford
- Hausman D (2012) *Preference, value, choice, and welfare*. Cambridge University Press, New York
- Infante G, Lecouteux G, Sugden R (2016) Preference purification and the inner rational agent: a critique of the conventional wisdom of behavioural welfare economics. *J Econ Methodol* 23:1–25
- Kahneman D (2003) A perspective on judgment and choice: mapping bounded rationality. *Am Psychol* 58:697–720
- Kahneman D (2011) *Thinking, fast and slow*. Farrar, Straus and Giroux, New York
- Kőszegi B, Rabin M (2007) Mistakes in choice-based welfare analysis. *Am Econ Rev* 97:477–481
- Le Grand J, New B (2015) *Government paternalism: Nanny State or helpful friend*. Princeton University Press, Princeton
- Loewenstein G (1996) Out of control: visceral influences on behaviour. *Organ Behav Hum Decis Process* 65:272–292
- Murakami K, Livingstone B (2015) Prevalence and characteristics of misreporting of energy intake in US adults: NHANES 2003–2012. *Br J Nutr* 114:1294–1303
- Salant Y, Rubinstein A (2008) (A, f) : choice with frames. *Rev Econ Stud* 75:1287–1296
- Sugden R (2008) Why incoherent preferences do not justify paternalism. *Const Polit Econ* 19:226–248
- Sugden R (2009) On nudging: a review of *nudge: improving decisions about health, wealth, and happiness*, by Richard H. Thaler and Cass R. Sunstein. *Int J Econ Busin* 16:365–373
- 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
- Tversky A, Kahneman D (1973) Availability: a heuristic for judging frequency and probability. *Cogn Psychol* 5:207–232