

Behavioral Economics: A Shaky Ground for Nudges

Annalisa Costella

1. Introduction

Ten years after its first publication, the excitement about “Nudge: Improving decisions about health, wealth, and happiness” (2009) has not yet diminished. If at all, it has increased. Methodological and normative critiques continue to be published both in philosophical and economic journals (Heilmann, 2014). *Nudge*, by behavioural economist Thaler and legal scholar Sunstein, proposes an approach to policy making that has come to be known as libertarian paternalism. Libertarian paternalism, as Sunstein (2018) puts it, is a paternalism of means, rather than ends. The idea is that people suffer from a number of biases when making decisions. Hence, they often take a suboptimal route to reach their goals. Libertarian paternalism enables people to satisfy those goals by guiding them towards the optimal route. Just like a GPS, libertarian paternalism offers people the best means to reach the ends that they themselves set out to achieve (ibid). It increases people’s *navigability* by offering them the best path to follow in order to satisfy *their* goals. To argue for their thesis, Thaler & Sunstein (2009) draw heavily on the findings of behavioural economics and related work in psychology. However, is it really the case that these findings lend empirical and conceptual support to the assumptions on which *Nudge* is based? This is the question that I will engage with in this essay.

I will argue that, contrary to what Thaler & Sunstein (2009) contend, the findings of behavioural economics alone do not provide sufficient theoretical support for nudges. To argue for it, I will first present the justification of libertarian paternalism that Sunstein and Thaler offer, by drawing on behavioural economics (Section 2). Subsequently, I will critically assess this justification. I will make use of related philosophical work on preference formation to argue that a prominent position in the field, the constructed preference view, does not warrant the interaction between preferences and choices that Thaler and Sunstein embrace. The *constructed preference view* only lends support to the existence of one category of preference mentioned by Sunstein (2018) (i.e. endogenous preferences). However, this category of preferences is problematic since it is incompatible with Sunstein and Thaler’s justification of nudges. Moreover, even if Sunstein and Thaler’s picture of the interaction between preferences and choices were correct, it is still the case that findings in behavioural economics do not provide sufficient support for the theoretical assumptions behind *Nudge*. Indeed, even if people had underlying true preferences, it would be hard to clearly distinguish them from the framing of those preferences (Section 3). I will then conclude with a plea for caution in embracing the leap from behavioural economic findings to nudges (Section 4).

2. Behavioural Economics as a justification for Nudge

In the lecture he gave when he received the Nobel Prize, Thaler (2018) referred to “Judgment under Uncertainty: Heuristics and Biases” (1974) by psychologists Kahneman and Tversky as his first discovery of a scientific approach that tackled what he had been observing for a while, namely people’s departure from the predictions of rational choice theory. With “rational choice theory”, I refer to a “representative individual” who has complete,

consistent and continuous preferences¹ The paper by Tversky & Kahneman, which started what would have later been known as the "Heuristics and Biases" program, pointed at the fact that people seem to use heuristics, which are rules of thumbs, when making decisions. Heuristics, in turn, might induce individuals to make predictable "mistakes" (Thaler, 2018). The idea expressed in that paper, and fostered by later works by Kahneman and Tversky (Kahneman, 2003) and a number of other authors (Rabin & Kőszegi, 2007; Camerer & Loewenstein, 2004), is that real-life agents act and behave very differently from rational agents. Individuals in real life do not have complete and continuous preferences. They overweight low probabilities and underweight high ones (Heilmann, 2014). They suffer from self-control problems, which make them display an inconsistent pattern of preferences (Thaler, 2018).

Let me provide an overview of the main findings of behavioural economics, before turning to how Sunstein and Thaler use them as a theoretical basis for justifying nudges. The review will be admittedly brief and incomplete, but it will shed some light on the intuition behind *Nudge*.

The body of work in behavioural economics that I will focus on is the one by psychologists Kahneman and Tversky. I will concentrate specifically on two major strands of their work, one on people's heuristics and biases, and another on framing effects and individual rationality. With regard to the first project, the main hunch behind Kahneman and Tversky's work is that people have two different modes of thinking and making decisions: System 1 and System 2 (Kahneman, 2003). System 1 is a cognitive process that is automatic, fast and emotionally charged (Kahneman, 2003). It usually draws on habits and heuristics (ibid.). System 2, in contrast, is slower, effortful and controlled. Mental signals are assigned to System 1 or 2 depending on the mental effort needed to handle them (ibid.). Hence, System 2 processes highly demanding mental tasks, while System 1 handles effortless ones. Since people's ability for mental effort is limited, they are usually unable to deal with many different effortful tasks. Contrary to effortless ones, demanding tasks disrupt each other. In this sense, Kahneman and Tversky have argued that System 2 "monitors" System 1. This is a shorthand for saying that individuals are more likely to react to tasks in a way that is not thought-through if they are already engaged in a demanding activity (Kahneman, 2003).

The idea of System 1 and System 2 is closely related to the concept of Planner and Doer that Thaler initially developed with Shrefin (Shefrin & Thaler, 1988) and that Thaler and Sunstein use recurrently in later work (Thaler, 2018; Sunstein, 2018). On Thaler's view, it is possible to build a two-self model of decision-making. According to this model, individuals' attitude to decision-making can be influenced by two contrasting dispositions, a "farsighted" planner and a "myopic" doer. The planner is a rational agent who tries to maximise her utility over time, while the doer easily indulges in short-term passions that contrast with the project of long-term maximisation. Thaler leaves the specific terms in which the planner and doer interact almost unspecified. However, he gives a hint by depicting them behaving in a principal-agent relationship, in which the principal (planner) attempts to induce the agent (doer) to do what he wants by either adopting pre-commitment strategies or relying on feelings of guilt developed after impulsive acts.

The second line of research pursued by Kahneman and Tversky regards framing effects and individual rationality (Kahneman, 2003). The idea is that, contrary to what economic theory predicts, the decision that an individual makes, given the set of options she is confronted with, varies depending on how the options at hand are

¹ In this essay, I refer to rational choice theory as the theory of choice presented in mainstream, widely used, economic textbooks such as Mas-Colell, Whinston, and Green, 1995 and Kreps, 2013. However, it should be noted that equating rational choice theory to the theory of choice presented in mainstream advanced economic textbooks is not consensually accepted in the literature (see, for instance, Gilboa et al., 2009). Moreover, this is only one view on rational choice theory.

presented (Kahneman, 2003). Hence, framing the same set of options in different ways influences (sometimes predictably) how the individual will choose (Kahneman, 2003). The idea of framing effects is also present in Bernheim's work in behavioural economics. According to him (Bernheim, 2016), it is possible to better understand people's evaluations of the options in a set if these are divided into direct and indirect judgements. Direct judgements are the judgements that people have over a set of alternatives, regardless of how those alternatives are presented. As Bernheim (2016, 17) puts it, direct judgements are judgements about "outcomes we care about for their own sake – our "ultimate objectives". Indirect judgements, in contrast, are judgements over the options that lead to the outcomes we care about. To make the difference clear, think of the following. I like the crunchiness of apples more than the softness of pears. This is a direct judgement. I prefer apples over pears. Imagine that I do grocery shopping online on a Dutch website. Since I do not understand Dutch, I need to translate the names of the different fruits to my native language. Hence, clicking on the fruit option involves an indirect judgement (i.e. my translation). If I believe, mistakenly, that "peer", in Dutch, means apple, then I will select some pears in my online shopping rather than some apples. My ending up with a worse option (pears, in this case) would be the result of a faulty indirect judgement. Behavioural economics, Bernheim argues, does not dispute my direct judgement of apples being better than pears. It only disputes the indirect judgement that leads me to buy pears rather than apples. By changing the framing of the options (i.e. in the example at hand by, for instance, attaching pictures of the different fruits to their respective names), a nudge can intervene on my indirect judgement in a way that makes it easier for me to choose the option that reflects my direct judgement.

Behavioural economics has produced a striking body of results that seems to be pointing at the following. People suffer from "biases" when making decisions. Hence, they might end up choosing an option or end up making a decision that is suboptimal for them. Therefore, changing the choice architecture, namely the environment in which people make a decision, without limiting the options that people have, can help them make better choices, *as judged by themselves* (Guala & Mittone, 2015). "Better choices" and "as judged by themselves" are both vague terms. Thaler and Sunstein are not explicit about their. However, they loosely use this condition to imply that individuals can evaluate a choice as better than another, upon reflection. "Better choices", in this context, stand for choices that better reflect people's preferences. If the satisfaction of preferences is taken to reflect an individual's well-being, as Thaler and Sunstein assume (Sunstein, 2017), then "better choices" has a further meaning. It means that those choices are better conducive to people's well-being. Nudges help exactly in this respect, according to Sunstein and Thaler. They are a tool for policy interventions, which help individuals make better decisions, as judged by themselves. Broadly understood, they steer people away from mistakes in their decision-making process by either eliminating biases or using those biases to trigger better decisions (Guala & Mittone, 2015).

This is just a brief and incomplete list of "biases" that people suffer from (for an extended one, see Kahneman (2003) and Thaler (2018)). Critically analysing the main findings of behavioural economics, however, is not the aim of the present essay. What I would like to present in the following subsections is, instead, how these findings have been used by Thaler and Sunstein to advocate for the use of nudges as policy interventions.

2.1. Nudges: from retirement to surgery

There are many examples of nudges implemented both at a governmental and firm level. Let me sketch a few examples:

Retirement plan: Barbara has problems in saving for retirement. She does not behave according to

the idea of life-cycle models of savings. According to the life-cycle hypothesis, Barbara should smooth her consumption over her life-time wealth. This means that, in every instance of her life, she should spend a constant fraction of her life-time wealth (Shefrin & Thaler, 1988). However, she does not do this. She is tempted by a number of spending opportunities which make her consume a higher portion of income than she would have liked to. If she were automatically enrolled in a retirement plan in the firm she works for, she would save a higher portion of her income. She would not bother to switch to the manual enrolment, and she would be grateful for it.

Healthy food in a cafeteria: Chris is on a diet. Hence, he has a long-term preference for healthy over unhealthy food, and should act consistently on this preference, according to the prescriptions of standard economics. However, every time he reaches the till of a cafeteria, he cannot refrain from buying his usual Mars bar that is placed right next to the till. It is just too tempting! Paul, the owner of the cafeteria, wants to help Chris in avoiding succumbing to temptation. He removes the Mars bar from its position and places it in an almost hidden position in the cafe. Chris is not tempted by the Mars bar anymore and can switch to a healthy apple instead. He is grateful for this.

Drinking campaign: Ellen is a student at Cornell. On a night out, she prefers sipping a glass of wine to downing the notorious “fishbowl”, a mixture of highly alcoholic liquors and sugary powders. If she were a rational agent, she would act consistently on that wine preference in every situation. However, everyone around her boasts about blacking out in the weekend and she feels socially pressured to do the same. As a result, she sips wine when she is out with close friends, but chugs a “fishbowl” at Frat parties, when she is surrounded by strangers. She acts inconsistently, since she *reverses* her preferences depending on the situation at hand. She happens to read an article on the university magazine where she discovers that only a low percentage of students in the whole university drinks so much to black out every weekend. She feels relieved and switches to her beloved glass of wine.

Surgery: Michael is afraid of surgeries and does not want to undertake a minor one. His doctor, whom he trusts fully, tells him that this surgery will change his life for the better, and convinces him of taking it. Michael undergoes the surgery and is grateful for this.

These are only some of the many possible examples of nudges. For an expanded list, see Thaler & Sunstein (2009). For the purposes of this essay, however, those examples are sufficient to identify nudges that target different *categories* of preferences.

2.2. *Three types of preferences as a target for nudges*

There are three types of preferences that nudges can apply to, according to Sunstein (2018). Those are:

Antecedent or context-independent preferences:

Individuals have clear antecedent preferences that nudges help them satisfy (Sunstein, 2018). This is the case of Ellen in **drinking campaign**. She has a clear preference for not drinking too much. Nudges help her to satisfy this preference.

Underlying true preferences:

Individuals might suffer from problems of self-control. They might have two different orders of preferences. A doer's preference for indulging to temptation and a planner's context-independent

preference for not doing so. This is the case of Barbara in **Retirement plan** and Chris in **Healthy food in a cafeteria**.

Context-dependent, endogenous preferences:

Individuals might not have stable, context-independent preferences. Their preferences might depend primarily on the context in which they get formed. Hence, their preferences are merely the product of the nudge. This is arguably the case of Michael in **Surgery**. His preference for undergoing or refraining from the surgery is merely the result of contextual cues. With a different doctor, he might have developed a different attitude towards the surgery.

So far, I have shown how nudges can address different categories of preferences, which I grouped in three main types, namely context-independent preferences, underlying true preferences and context-dependent ones. With this framework in mind, I will set out to investigate whether the justification that is standardly provided for nudges can hold for all these types of preferences. I argue that findings in behavioural economics can justify nudges if two conditions are jointly satisfied. On the one hand, it should be the case that work in behavioural economics and psychology lends support to the existence of (at least one of) these three categories of preferences. On the other, it should hold that the existing type(s) of preference, by itself (themselves), is (are) compatible with the justification for nudges provided by behavioural economists, which is that nudges only help people taking the optimal route to satisfy *their* preferences, as judged by themselves. This means that there should be quite robust evidence that either:

- a) People have antecedent, context-independent preferences, which nudges help satisfy.
- b) People have underlying true preferences, which they fail to satisfy because of weakness of will. Nudges help them satisfy these preferences.
- c) There are instances in which people have context-dependent preferences. Nudges frame the environment in which people make decisions in a way that helps them to form preferences that they deem desirable to retain.

It is worth noticing that the *joint* satisfaction of a), b) and c) is not necessary to justify nudges. It might be the case that only one of these conditions holds. This would be enough to justify nudges in the relevant respective scenario. However, the scope of nudges would be significantly diminished.

Before analysing whether a), b) and c) are satisfied, let me focus on the second condition, which is whether the standard justification for nudges in behavioural economics holds valid for these three types of preference. The two first categories of preferences, if existing, supposedly form a quite straightforward theoretical basis for the nudge project. The idea in these two cases is that people have true, underlying preferences that exist in their head and that they are able to access (imperfectly, as in the cases formulated above) (Bernheim, 2016). By reframing the environment in a way that better enables people to make choices that conform with their true, underlying preferences, nudges would simply help them satisfy those preferences. The third category is more controversial, since it is unclear how nudges can help people in making "better choices" when people do not have clear true preferences for a certain outcome. Those cases, as Sunstein (2018) himself admits, are more complex. If there are no true, underlying preferences, then it becomes unclear how nudges help people to reframe the choice architecture in order to make those people better off.

Time to take stock. Behavioural economics has produced a striking amount of evidence that people suffer from mistakes and biases when making decisions. Those mistakes and biases prevent them from making choices that are optimal, according to their own judgement. Nudges are policy interventions that steer people away from those mistakes, by addressing three main categories of preferences. These are context-independent preferences, underlying true preferences and endogenous preferences. I have argued that, if the first two exist, then the justification for nudges provided by behavioural economists such as Sunstein and Thaler holds valid. In contrast, this justification for nudges is not sufficient if only the third category exists.

3. Behavioural Economics as a justification for Nudge?

At a first glance, Thaler and Sunstein's argument seems compelling. But does it hold up to scrutiny? I believe it does not, and I will argue for this in the remainder.

My argument proceeds in two steps. First, I argue that prominent work on preference formation in philosophy of science does not lend support to the existence of the first two categories of preferences mentioned in Section 2. Second, I argue that, even if those findings were incorrect, there is still a theoretical difficulty. Disentangling preferences from their framing is difficult and, consequently, the identification of preferences is not a straightforward task. As a result, what counts as a preference and what counts as framing might partly be the artefact of the behavioural economist herself. Hence, the justification for nudges provided by Thaler and Sunstein, which is that they help people better satisfy their preferences, as judged by themselves, does not hold.

Let me start with the first step. One prominent view in philosophy of science about how preferences get formed is the *constructed preferences view*, which is backed by recent psychological findings (Bernheim, 2016). According to it, individuals do not have a true, underlying preference for a specific outcome. They form a preference only in the moment in which they are asked to make a choice. In that specific moment, they aggregate a number of aspects of their current experience and form a preference that did not exist before. As Bernheim (2016, 20) puts it, "from this perspective, the concepts of "true preferences" and "experienced utility" are fictions; they may play useful roles in "as-if" representations of behaviour, but we should not take them literally." The view according to which it is implausible that people have underlying true preferences is expressed also by Guala & Mittone (2015) and Sunstein (2018).² According to both, it is psychologically questionable to think of choices as the result of an interaction between underlying true preferences and biases in the decision-making process. If the constructed preference view is correct, then neither the first category of preferences (antecedent preferences) nor the second one (underlying preferences) are grounded on a valid philosophical account of preference formation (that is psychologically informed). In contrast, the third category, namely the one of context-dependent endogenous preferences, does not seem incompatible with the constructed preference view. Indeed, if people form preferences only when they are called upon to do so, and those preferences are highly dependent on the framing, then it is plausible to assume that their preferences are context-dependent and endogenous. However, this third category is also the most problematic for nudges. The reason is that nudges are presented as tools for policy making that change the framing of

² For related psychological work on the subject, see Simon, Krawczyk, Bleicher, and Holyoak, (2004, 2008); Lichtenstein and Slovic, (2006).

the options in a way that enables people to better satisfy their preferences. Nevertheless, if people's preferences are the mere result of framing, then it is unclear why changing from one framing to another should enable people to better satisfy their preferences, if these do not exist outside the framing itself. Hence, it is debatable whether the existence of context-dependence preferences justifies nudges at all. If context independent preferences do not exist and endogenous preferences do not point to any specific set of preferences that an individual has reason to uphold, the findings of behavioural economics *alone* are not sufficient to justify nudging in the way that Thaler and Sunstein do. That is, nudges cannot be defended on the sole ground that they are tools that help people better satisfy their own preferences.

A supporter of nudges, however, could object in two ways. She could argue that the constructed preference view, even if prominent among philosophers of science, is only one among many possible accounts of how preferences are formed in psychology³ Hence, its validity is not fixed in stone and it might turn out to be a wrong theory. Or she could point at psychological literature that shares the view that preferences are constructed but argues that stable preferences are not incompatible with this view⁴ Let us assume that either of these is the case. Hence, either preferences do not form in the way that the "constructed view" envisions or, even if they do, they can be stable over time. Nudges could then be limited to addressing either the first two categories of preferences or preferences that are formed on the spot but are stable over time. Would this clear the ground for a behavioural economic foundation of nudges? No, according to Infante et al. (2016a,b). This leads me to the second step of my argument.

Imagine, as Infante et al. (2016a) ask us to do, that SuperReasoner, an individual who has underlying true preferences and whose judgement is governed exclusively by System 2, the slow, forgetful reasoning mode explained in Section 2. She has no cognitive limits. She possesses infinite information. How would she behave in situations such as **Retirement plan** and **Healthy food in a cafeteria**? Differently from Barbara and Chris? Infante et al. argue that she would not. What they argue is that SuperReasoner in the cafeteria might well choose the Mars bar if this is close to the till and, the following day, the healthy snack if this is the one right next to the till. This choice is inconsistent only under the reading that SuperReasoner can have exclusively a context-independent preference for healthy snack over Mars bar rather than a preference for "healthy snack next to the till" over "healthy snack far away from the till". Her preferences are inconsistent only under the reading that the only correct way to interpret the distance between the food and the till is merely a matter of framing, rather than being part of the preference itself. In turn, the supporter of nudges might be highly sceptical of this example. Indeed, she might object that in the cafeteria case it is hard to argue against the stance that the position of the food is clearly only a matter of framing. The direct judgement of SuperReasoner is obviously the one between the healthy snack and the Mars bar. The remainder (i.e. the position of the food) is plainly the result of indirect judgement. This might well be the case, but it does not say anything about the generalisability of our nudges' supporter latter remark, namely that framing and preferences can be disentangled straightforwardly in real-life scenarios. Indeed, there are many real-life situations in which clearly distinguishing between the framing and the underlying preference that a person has might be extremely complex. To see it, think of **Retirement plan**. How should one distinguish the underlying preference for consuming or saving from its framing? Hence, even if people had true underlying or

³ See, for instance, Caleb Warren, McGraw, and Van Boven, (2011).

⁴ See, for instance, Bettman, Luce, and Payne, (2008) for a similar point.

context-independent preferences, disentangling those preferences from their framing is a highly complex task. The reason is that the behavioural economist would need to understand what the chooser herself considers the object of her preferences and what its framing. This consideration not only changes from one person to another but might be not so clear even to the chooser herself. If true preferences and framing of those preferences cannot be distinguished, it is unclear how nudges could intervene exclusively on framing effects, when those have not been clearly individuated. Hence, if true, the findings of behavioural economics, on their own, do not present a sufficient justification for policy interventions such as nudges.

In a nutshell, I have argued that Behavioural Economics does not support nudges on its own. There are two main reasons for this. Firstly, the constructed preference view only provides support for the existence of context-dependent, endogenous preferences. Since endogenous preferences are exclusively the result of framing, and nudges change the framing from one to another, it is hard to see how this change in the choice architecture can help people better satisfy preferences that they do not have in the first place. This first step in my argument leaves open the possibility that, if the constructed preference view were incorrect, then nudges could still be justified in cases in which people have underlying true preferences. I rule out this concern by arguing that there is a stronger conceptual reason for why findings from behavioural economics alone cannot provide a sufficient justification for nudges. This is that, even if people had underlying true preferences, distinguishing them from the framing is difficult. In many instances, deciding what counts as framing and what as preference is merely the result of an arbitrary artifice carried out by the behavioural economist.

4. Conclusion

I have argued that findings from behavioural economics *alone* are not enough to provide a justification for the theoretical assumptions behind *Nudge*. To argue for this, I have first laid out Thaler and Sunstein's argument: that findings from behavioural economics provide the theoretical basis on which the nudge program is founded. Subsequently, I have analysed this claim. I have argued that it is problematic for two reasons. First, not all the three categories of preferences that Sunstein (2018) identifies seem to be compatible with a prominent model of preference formation. Indeed, only one of them, endogenous preferences, would be compatible with this model. However, Thaler and Sunstein's justification for nudges does not hold up to scrutiny if the only existing preferences are endogenous ones. Second, even if this first objection did not hold, there is still a stronger, conceptual objection. That is, even if the first two categories of preferences highlighted by Sunstein were compatible with a psychological model of preference formation, behavioural economic findings *alone* would not be able to provide sufficient justification for the theoretical assumptions behind nudges. The reason for it is that the distinction between underlying preferences and their framing is often the artefact of the behavioural economist herself.

Acknowledgements

I would like to thank Jack Vromen for the sharp feedback on an initial version of this essay and for nominating it to the editors of ESJP. By doing so, he provided me with my very last opportunity to engage with the extremely stimulating activity of contributing to a very well-run student journal. Moreover, I am grateful to some anonymous ESJP referees. Their patient and constructive feedback on

a number of previous versions of this essay challenged me to rethink and revise my essay extensively.

Annalisa Costella (1995) obtained a BSc in Economics at Bocconi University in Milan, where she was born and raised. She then moved to the London School of Economics to attend an MSc in Philosophy of the Social Sciences. She continued her education with a Research Master in Philosophy and Economics at Erasmus University, where she is currently doing her PhD in Philosophy. Her life in Rotterdam did not change substantively from the previous edition. She still complains about Dutch cooking skills, and is still equally unable to play tennis. The only addition is her new willingness to learn how to climb, an attempt that is likely to end up in the same way as tennis did.

References

- Bernheim, D. 2009. "Behavioral welfare economics". *Journal of the European Economic Association*, 7 (2-3): 267–319.
- Bernheim, B. Douglas. 2016. "The Good, the Bad, and the Ugly: A Unified Approach to Behavioral Welfare Economics." *Journal of Benefit-Cost Analysis* 7 (1): 12–68.
- Bettman, J. R., Luce, M. F., and Payne, J. W. 2008. "Preference construction and preference stability: Putting the pillow to rest." *Journal of Consumer Psychology*, 18 (3): 170–174.
- Camerer, C., and Loewenstein, G. 2003. "Behavioral economics: Past, present, future." In *Advances in Behavioral Economics*, 3-51.
- Gilboa, I., Postlewaite, A., and Schmeidler, D. 2009. "Is it always rational to satisfy Savage's axioms?" *Economics & Philosophy*, 25 (3): 285-296.
- Guala, F., and Mittone, L. 2015. "A Political Justification of Nudging." *Review of Philosophy and Psychology*, 6 (3): 385–395.
- Heilmann, C. 2014. "Success conditions for nudges: A methodological critique of libertarian paternalism". *European Journal for Philosophy of Science*, 4 (1): 75–94.
- Infante, G., Lecouteux, G., and Sugden, R. 2016a. "Preference purification and the inner rational agent: a critique of the conventional wisdom of behavioural welfare economics." *Journal of Economic Methodology*, 23 (1): 1–25.
- Infante, G., Lecouteux, G., and Sugden, R. 2016b. "On the Econ within: a reply to Daniel Hausman." *Journal of Economic Methodology*, 23 (1): 33–37.
- Kahneman, D. 2003. "Maps of Bounded Rationality: Psychology for Behavioral Economics." *American Economic Review*, 93 (5): 1449–1475.
- Kreps, D. M. 2013. *Microeconomic foundations I: choice and competitive markets (Vol. 1)*. Princeton university press.
- Lichtenstein, S., and Slovic, P. 2006. *The construction of preferences*. Cambridge University Press.
- Mas-Colell, A., Whinston, M., and Green, J. 1995. *Microeconomic theory*. New York: Oxford University Press.
- Rabin, M., and K szegi, B. 2007. "Mistakes in Choice-Based Welfare Analysis." *American Economic Review*, 97 (2): 477–481.
- Shefrin, H. M., and Thaler, R. H. 1988. "THE BEHAVIORAL LIFE-CYCLE HYPOTHESIS." *Economic Inquiry*, 26 (4): 609–643.
- Sunstein, C. R. 2018. "Better off, as judged by themselves: a comment on evaluating nudges." *International Review of Economics*, 65 (1): 1–8.
- Thaler, R. H. 2018. "From Cashews to Nudges: The Evolution of Behavioral Economics." *American Economic Review*, 108 (6): 1265-87.

- Thaler, R. H., and Sunstein, C. R. 2009. *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.
- Tversky, A., and Kahneman, D. 1974. "Judgment under uncertainty: Heuristics and biases." *Science*, 185 (4157): 1124–1131.
- Simon, D., Krawczyk, D. C., Bleicher, A., and Holyoak, K. J. 2008. „The transience of constructed preferences." *Journal of Behavioral Decision Making*, 21 (1): 1–14.
- Warren, C., McGraw, P., and Van Boven, L. 2011. "Values and preferences: defining preference construction." *Wiley Interdisciplinary Reviews: Cognitive Science* 2 (2): 193-205.

