

Nudge plus: incorporating reflection into behavioural public policy

Sanchayan Banerjee and Peter John

January 2020

**Grantham Research Institute on
Climate Change and the Environment
Working Paper No. 332**

ISSN 2515-5717 (Online)

The Grantham Research Institute on Climate Change and the Environment was established by the London School of Economics and Political Science in 2008 to bring together international expertise on economics, finance, geography, the environment, international development and political economy to create a world-leading centre for policy-relevant research and training. The Institute is funded by the Grantham Foundation for the Protection of the Environment and a number of other sources. It has six research themes:

1. Sustainable development
2. Finance, investment and insurance
3. Changing behaviours
4. Growth and innovation
5. Policy design and evaluation
6. Governance and legislation

More information: www.lse.ac.uk/GranthamInstitute

Suggested citation:

Bannerjee S and John P (2020) *Nudge plus: incorporating reflection into behavioural public policy*. Grantham Research Institute on Climate Change and the Environment Working Paper 332. London: London School of Economics and Political Science

This working paper is intended to stimulate discussion within the research community and among users of research, and its content may have been submitted for publication in academic journals. It has been reviewed by at least one internal referee before publication. The views expressed in this paper represent those of the author[s] and do not necessarily represent those of the host institutions or funders.

Nudge plus: incorporating reflection into behavioural public policy*

Sanchayan Banerjee⁺ and Peter John[^]

⁺PhD Candidate in Environmental Economics, London School of Economics,
S.Banerjee9@lse.ac.uk

[^]Department of Political Economy, King's College London, peter.john@kcl.ac.uk

13 January 2020

*We thank other members of the project team—Susana Mourato and Matteo Galizzi—and we are grateful to LSE for support: Sanchayan's research has benefitted from the LSE Robert and Dilys Rawson Scholarship supporting his PhD work and support from the Grantham Foundation for the Protection of the Environment. We also thank those who commented on an earlier draft of this paper: Manu Savani, Sarah Cotterill, Anomitra Chatterjee, and Dario Krpan. We thank the Centre for European Studies at Science Po, Paris, for a seminar on 26 November 2019 where we presented the paper, in particular Charlotte Halperin for discussant comments. We are grateful to Till Grune-Yanoff for participating in an informal seminar on nudge plus and boost which took place at the LSE on 28 November 2019 and for his comments on our paper.

Abstract

We outline a modified version of behaviour change called nudge plus, which incorporates an element of reflection as part of the delivery of a nudge. Nudge plus builds on recent work advocating educative nudges and boosts. Its argument turns on seminal work on dual systems that presents a more subtle relationship between fast and slow thinking than is commonly assumed in the classic literature in behavioural public policy. Our claim is that a hybrid nudge-think strategy can be a useful additional way to design pro-social interventions. We review classic and recent work on dual systems to show that a hybrid dual process account is more plausible than the default interventionist or parallel competitive framework. We put forward a way to operationalise nudge plus and set out what reflection could embody. We compare nudge, nudge plus, and boost, and draw testable implications.

Key Words, *nudge; boost; nudge plus; dual process theory*

A nudge that incorporates an element of reflection might first seem to be a contradiction in terms. After all, the whole point of a nudge is that it happens automatically without much conscious thought on the part of the individual. The acknowledgement of the low cognitive capacity of individuals to make fully rational choices is thought to be one of the key advantages of a nudge over other policy instruments, such as information campaigns, laws, or taxes (Thaler and Sunstein 2009). Nudge is supposed to work on fast and automatic system 1 processes in a way that leaves the slow and reflective system 2 unengaged. Thaler and Sunstein have been at pains to stress that individuals could reflect and agree with a nudge after the event rather than before or during it, hence their commitment to publicity (Thaler and Sunstein 2009, 244). Moreover, Sunstein reports survey evidence that suggests that citizens approve of nudges being done to them of which they might not be fully aware (Sunstein 2016b, 140–41).

Nonetheless, a modified account of behaviour change is starting to emerge in a number of recent contributions, which make the claim that a nudge might work better and more legitimately if it incorporated an element of self-awareness and internal deliberation. In particular, a nudge based on some degree of reflection might be more capable of generating long-term, persistent, and sustainable behavioural changes than classic nudges can achieve alone. For example, Mühlböck et al. (2019) show that adding a reflective survey before a classic information nudge increases the uptake of the information nudge and reduces the unemployment period among Austrian youth. Recipients of a nudge welcome sequential additions of reflective strategies.

It may already be the case that many existing nudges already have an element of self-reflection already built into their delivery, which could be acknowledged and enhanced. A commitment device, for example, are based on the idea that a pre-commitment default keeps people to a desired course of behaviour; yet ensuring an

individual has some autonomy and reflection to think through the nature of what is involved, which precedes the signature of the commitment contract (Stutzer, Goette, and Zehnder 2011). The concern with reflection also appears in Sunstein's advocacy of educative nudges and his contrast between system 1 and system 2 nudges (Sunstein 2016a). Some classic nudges can be made to work better by simultaneously bundling them with reflective elements to make them more salient to the receiver, such as over-willingness to pay for flood risk insurance (Bradt 2019), encouraging job seekers to think more slowly (Heller et al. 2017), and training for self-investment of Liberian youth (Blattman, Jamison, and Sheridan 2017). The idea that the conscious brain may be engaged with nudges appears in the recent debate about whether being aware of a nudge affects its efficacy, with an answer from the lab that it does not (Loewenstein et al. 2015).

To move the debate forward, John and Stoker set out a new concept called 'nudge plus', which adapts their earlier contrast between nudge and the deliberative 'think' (John 2018; John and Stoker 2019). Nudge plus refers to interventions that might have both nudge and think incorporated into them. A nudge plus can have a simple modification, such as a reflection tool as part of the opt-out default, that stresses the salience of the agent. The agent receiving a default combined with a plus would switch from thinking fast to thinking slow in a way that would help responding to the classic nudge. The combination of the classic nudge with a reflective plus is nudge plus.

However, nudge plus requires careful justification. It must rest on a coherent and defensible account of cognition that also makes sense in terms of behavioural public policy to ensure it is operationalizable. It is essential to understand how both nudge and nudge plus relate to the dual process models as advanced by Stanovich and West (2000), then taken up by Kahneman (2012) and Thaler and Sunstein (2009).

Crucially, it is important to defend, using theory and recent research in psychology, how system 1 and 2 processes can be in play at the same time. Moreover, it is not clear how nudge plus works best when applied; in particular is the nudge and plus combined into the same intervention, or whether the plus precedes the nudge or vice versa? Would it be best, as Hertwig and Grüne-Yanoff (2017) argue, to boost individual capacity for decision-making? All these questions need to be resolved when setting out a research agenda for behavioural public policy incorporating reflection, so it coheres as a research and policy programme. This is what this paper sets out to do.

To outline a modified version of behaviour change, this paper proceeds first by defining its terms and reviewing the core literatures—that of classic nudge and boost. The key part of the paper is a review of the literature from cognitive science that supports the engagement of the two systems and questions the idea that human decision and action can be explained solely by a singular type of brain processes as the default. A more practical part of the paper offers clarity on what reflection means and its purpose. It then moves on to discuss the sequence of nudges and reflective exercises. We use these insights to compare nudge, nudge plus, and boost across common dimensions; then set out some testable propositions. We conclude by making some claims about the future of nudge plus relative to other behavioural tools.

Classic nudges

As set out by Thaler and Sunstein (2009), nudges are low cost signals or procedures that encourage, from the planner's point of view, a socially desirable change in behaviour in ways that preserve the autonomy of the individual without restricting the choice set. They draw on about twenty-five years of research in behavioural economics that locates the origins of human behaviour in psychological processes and modifies a simple rational-cost calculation through the prism of the heuristics and biases (see also Sunstein and Thaler 2003).

Their definition of nudge, much cited, is this: 'nudge, as we will use the term, is any aspect of the choice architecture that alters people's behaviour in a predictable way without forbidding any options or significantly changing their economic incentives. To count as a mere nudge, the intervention must be easy and cheap to avoid. Nudges are not mandates. Putting fruit at eye level counts as a nudge. Banning junk food does not.' (Thaler and Sunstein 2009, 6). Examples of nudges include pension defaults, organ donation forced choices, and commitment devices for diet. They require changing the choice architecture of rules and procedures that, in part, governments and other agencies control from their command of bureaucratic and legal processes that affect choices of citizens. Although there is a considerable debate about the definition of a nudge (see Baldwin 2014; Oliver 2017), we refer to an instrument involving 'changes in choice architecture' only. Information and labelling fall beyond the purview of the classic nudge.

Boost

A boost is a class of behavioural policies that seek to improve the decision-making power or competence of an agent. It is different from the other behavioural instruments as it is directed to increase cognitive capacity only. As Hertwig and Grüne-Yanoff put it, the goal of boosts is to ‘improve people’s competence to make their own choices and to make it easier for people to exercise their own agency by fostering existing competences or instilling new ones’ (2017, 2). A boost is based on an alternative psychological theory to reason why humans depart from fully rational behaviour and show how to improve their decision-making process by upgrading their ‘repertoire of decision-making skills’ (the adaptive toolbox). A boost goes beyond regular schooling mechanisms. A good example of a boost is an uncertainty management rule to interpret advice given to health patients as to how to make good choices, such as over treatments that might vary in the likely outcomes. People find it very hard to understand these probabilities, as Kahneman and Tversky (1982) found in their experiments. With some training, individual capacity can be increased to make better decisions.

Rather than being just another device to improve rationality, a boost is based on a different assessment of cognition than nudges: people are intuitive and frugal in their use of the minds, and that interventions need to be targeted to make best use of the common sense that people have innately. With a bit of guidance, people can be taught to be Bayesians, for example. The boost works within the prism of the simple heuristics approach as agents, given their cognitive burden, choose shortcuts which are often cognitively beneficial but could go wrong at times. However, such biases are not systematic and certainly not always tied to the heuristics. Instead of getting rid of all heuristics, this approach believes in making such heuristics smarter and intuitive to

avoid those occasional mistakes. Heuristics work best when the agent's cognitive skillset and the external environment work in tandem.

The boost, however, assumes that agents have the motivation and competence to benefit from the improved decision-making processes resulting from the boosting mechanism. This is different to the classic nudge which assumes a 'somewhat mindless, passive decision maker' (Thaler and Sunstein 2009, 36). A boost assumes that by changing the environment, or their cognitive skills and abilities (competency), an agent will make better decisions (Grüne-Yanoff and Hertwig 2016). Different variants of these boosts have been tested in varied settings; for instance, literacy skills have been shown to improve financial decision making (Drexler, Fischer, and Schoar 2014), and quick dietary rules have helped people make more healthy food choices (Pollan 2010).

Nudge plus

One critique made by researchers outside behavioural economics is that nudges can only deal with relatively minor public problems and ones that are strictly under the guidance of the benevolent policy-maker, using techniques of behavioural science (e.g. Mols et al. 2015; Marteau et al. 2011). The size of the challenge achieving sustainable behaviour change requires a more profound and long-lasting solution that builds on the consent of individuals. There also needs to be a way to address the criticism that nudges manipulate so reduce autonomy of individuals or even bypass their explicit consent, even from within a libertarian paternalist framework (Bovens 2008). These objections to nudge are addressed in an alternative programme of 'think' that implies that debate and deliberation can help individuals achieve their objectives and tested in a series of interventions (John et al. 2013; 2019). But thinks are hard to scale up to the general

population; they involve much time and strong commitment from individuals that they may not wish to give. To be closer to the classic nudge, John and Stoker (John and Stoker 2019) set out a modified version of nudge—nudge plus—that incorporates an element of reflection, yet might be cognitively easy for an agent to uptake.

Nudge plus takes a different approach to cognition. While a classic nudge takes advantage of the biases of system 1 processes by co-opting it (internal architecture) and changing the external environment (choice architecture), a nudge plus builds on it by adding reflective strategies to the nudge. The reflective strategy, the plus, offers to extend the autonomy of the agent. Nudge plus can be thought to lie on a continuum of agency autonomy, with the left of the spectrum (no autonomy) being occupied by classic nudges and the right (complete autonomy) as pure reflective strategies. Nudge plus can be potentially more liberty preserving and transparent than the classic nudge.

What might a simple nudge plus look like? Building on the classic opt-out defaults, it might be easy to think of nudge plus as an opt-out default, say a duplex printing set-up, combined with a self-reflective commitment nudge that leads the agents into thinking about the purpose of their print and choosing the correct format thereafter. Similarly, a naïve thought experiment might indicate that, for instance, people who see fruit instead of chocolate near the checkout tills in a cafeteria, the classic Thaler and Sunstein nudge, might wish to have a chance to reflect on why is the fruit placed there as addiction to sugar will not cause them to substitute the fruit without some kind of thought (John 2018, 129). The prompt could be a debate about climate change with the cashier, or some kind of exercise that happens while customers are waiting, such as a push button survey, or if the customer had previously learnt of the NHS ‘five-portion-a-day’ rule. The plus could also happen before the nudge, such as at the start of the

queue, or even afterwards to encourage consistent consumption of the fruit on later occasions.

Practical as these examples appear, they involve a shift in the model of cognition assumed to be in operation when people are making choices. To resolve this issue, we present evidence from the cognitive psychology literature that justifies that a hybrid nudge-think framework as reasonable. The approach needs to be situated within a defensible account of dual system theory. The upcoming section will give a brief overview of the different dual process theories. We define what reflection actually means, what it aims to achieve and how could such a nudge plus be operationalised before providing a clearer exposition into the conceptual differences between the different accounts of behaviour change.

Dual process theories: an overview

Canonical work in social cognitive psychology bifurcate cognitive processes into two different kinds, laying the foundational stone of what later came to be broadly known as the dual process theories. Dual process theories posit that, ‘there are two distinct processing models available for cognitive tasks: one (type 1) that is fast, automatic and non-conscious, and another (type 2) that is slow, controlled and conscious’ (Frankish 2010, 914). These theories date back to the 1960s and have been evolving ever since. Although different schools of thoughts have emerged within the domain of the dual process theories, the central emphasis has been on the idea that cognitive processes can be clearly distinguished into an intuitive (or, heuristic) and analytical (or, systematic rule-based) type that might interact with one another, and take precedence over each other depending on the nature of the task. These theories came to existence

independently until an attempt was made to combine them into a more structured and generalised framework based on the common traits of these models (for a summary, see Gawronski and Creighton 2013, chapter 14).

A central tenet of these dual process theories is the clear distinction between the two types of brain processes. Frederick and Kahneman (2002) show how the cognitive processes labelled as system 1 and system 2 were popularly used following Stanovich (1999) and Stanovich and West (2000), then later popularised by Kahneman (2012) in his book *Thinking, Fast and Slow*. These systems have a set of characteristics: the more reflexive system 1 is usually thought to be automatic, effortless, associative, rapid, and parallel, and is often contextualised as the old mind whereas the reflective type system 2 is often summoned as the controlled, effortful, reasonable, lazy, and rule based. However, the labels given to the two class of processes are possibly misleading and imply two distinct types of cognitive processes interconnected by neurological traits only (for details, see Lieberman 2003).¹ Next we discuss the interplay between the two types of cognitive processes and the main critiques of dual process theories in recent scholarly literature.

Critiques of dual process accounts

Dual process models drew criticism from a range of social cognitive psychologists; ones that proposed a single, malleable and unified brain processes (Osman 2004; Kruglanski and Gigerenzer 2011) to ones that posited multiple processes models (Sherman 2006). While these dual processes were characterised by multiple traits, they

¹ Note that system 1 and 2 processes are interchangeably used with type 1 and 2 processes.

suffered conjointly from a clustering problem i.e. most of these traits failed to co-occur with the processes involved; for instance, as Bargh (1994) posited in his four horsemen of automaticity theory, the processes could be explained on basis on four attributes: awareness, intention, efficiency, and cognitive control. However, as it turned out, none of these were always concurrent in any type of cognitive processing involved; for instance, a process that was fast and inattentive turned out to be cognitively controllable, thereby failing to satisfy the label of being truly reflexive in nature.

More recently, Evans and Stanovich (2013) seek to resolve the debate by positing that the two set of processes share multiple features but they are not all defining. They put forward a necessary and sufficient condition for each of the two dual processes: Type 1 processes must have ‘autonomy’ and type 2 processes must satisfy ‘cognitive decoupling’ for hypothetical thinking. However, assertions regarding the existence of one, two, or multiple processes are ontological in nature. As Gawronski and Creighton put it, ‘we cannot test empirically if there are one, two or multiple processes. Researchers can make decisions about the usefulness of ontological claims by empirically testing assumptions about the proposed processes.

Dual process theories have fared very well and have a dominant role in social psychology (Gawronski and Creighton 2013, 307–8). However, there is still some disagreement over how the systems operate together (Zimmerman 2016). In line with recent neuroscientific evidence in social cognitive psychology, the two sets of neurological processes are interconnected by a common defining characteristic. In simpler terms, the brain has two types of processes, one that is autonomous and other that can think, and every action-decision pair is a result of these processes, which might either work simultaneously or sequentially.

The interplay of cognitive processes

A key issue is how these brain processes interact with one another. Does one take precedence over the other? Can both the processes occur concurrently? Can one type be put off while the other works? To resolve this conflict amongst the dual processes, Evans (2007) put forward a simple additively separable model of decision making under uncertainty. In this analysis, borrowing from Evans, it is possible to simplify the story with Kahneman's characters (slightly renamed): Bob/Bobbie (type 1) and Joe/Joanna (type 2). For brevity, let's assume Bobbie is the autonomous character, acts on her free will and is quick to answer things; while Joanna is capable of cognitive reasoning and decoupling.

Let us further assume that Bobbie and Joanna have participated in a pub quiz as a team. Bobbie and Joanna will get a minute to answer, and there are four thematic rounds. Bobbie and Joanna are both normatively rational.² How do they decide on their answers as a team? There could be two possibilities: either Bobbie and Joanna split the themes on the basis of their thematic expertise such that whoever fits the bill answers the question in that round, or both of them decide to answer in all the rounds. If the latter holds, they might have to write their individual answers on a chit and then cross check before reaching a conclusion. Let us suppose Bobbie and Joanna choose the former. This is what Evans (2007) classifies as the pre-emptive conflict resolution dual process models and resonates with the dual process theory put forward by Klaczynski (2000) and the selective scrutiny model of belief bias (Anonimo et al. 1993).

² A common fallacy of dual process theories is the presumption that only Joanna is normative and rational (for discussion, see J. Evans 2007; and J. Evans and Stanovich 2013, 7).

What happens if Bobbie and Joanna walk down the other path, and decide to answer collectively in all thematic rounds? There are two possibilities again; either, Bobbie and Joanna end up writing the same answer and there is no conflict, or, they write different answers and end up with a conflict. If both write the same answer, the solution is trivially the same. However, in case of a conflict, they must find a way to resolve the conflict together. Evans (2007) proposes a probabilistic solution to this: both Bobbie and Joanna try first, and one's opinion will be heard and finally settled upon after a verbal duel between the two. This is referred to as the parallel-competitive model of dual process theory. Contrarily, one might always dominate the other in case of a conflict as if Joanna wins a verbal duel whenever there is one. This conflict resolution model is referred to as the default-interventionist model of dual process theory. In this case, both Joanna and Bobbie can dominate one another (see A-dominating or C-dominating process dissociation model by Jacoby and Lindsay (1994) and Jacoby (1991) respectively). However, the classic default-interventionist models (J. Evans 2010; Kahneman 2012; J. Evans and Stanovich 2013) assume a pertinently corrective role for Joanna at all times i.e. the system 2 processes will override system 1 processes if there is a conflict detected.

Dual processes: simultaneous or sequential?

Even within the class of dual process theories, there can be a subcategorization depending on the type of the response mechanism involved. Although there are different interaction mechanisms of the dual processes, does one suit the role better than the other? This remained an ontological concern for a long time until recently, and 'all three models enjoy[ed] implicit support from dual process theorist' (Evans, p10, 2007);

for instance, Epstein's cognitive-experiential self-theory (1994) presume that the two types of processes (called systems incoherently) might occur in parallel, while Kahneman and Frederick (2002, 51) endorse a default-interventionist structure as they wrote, 'we assume, system 1 quickly proposes intuitive answers to judgment problems as they arise, and system 2 monitors the quality of these proposals, which it may endorse, correct or override'.

A common argument in favour of the default-interventionist mechanism has been that type 1 processes are fast and automatic, while type 2 processes are slower and rule-based. If they have to operate in parallel, why would the faster type 1 process wait for the slower type 2 processes. This has been countered in various ways in the literature; we will put forward two major arguments for this. First, these processes share common characteristics but not all of them might be defining and co-occur. What is being perceived as slow, rule-based and reasoning might also favour heuristics, and lead to anomalies. Second is an analogy that we borrow from Paap, Noel, and Johnson (1992) and noted by Evans (2007, 12, fn 4): it posits that in a horse race, but it is not always the fastest horse that wins. In a very similar fashion, the dual processes can run in parallel without Bobbie always ending up with an answer before Joanna.

However, DeNeys (2012; 2014; 2019) has suggested a hybrid dual process theory. It overcomes the shortcomings of the prior generic processing architectures that guide the interplay of the two cognitive processes. In explaining this hybrid model, Gronchi and Giovanelli (2018, 2) posit that a 'shallow analytic monitoring process is always active to detect potential conflicts between the two systems, and an optional deeper processing stage is activated once an actual conflict between fast and slow thinking is found'. To facilitate this hybrid cognitive architecture, DeNeys is suggestive of two types of system 1 responses; one that is heuristically driven (the so-called

intuitive processes) and the other which is logically intuitive. According to this hybrid model, the two system 1 processes are activated in parallel followed by the system 2 processes which act to validate and justify the conflict resolution put forward by the logically intuitive processes. In this novel attempt to bridge the gap between the default-interventionist and the parallel competitive model, DeNeys upgrades the role of system 1 processes; he acknowledges that system 2 processes is more of the validator than the corrector as has been incorrectly assumed by the prior scholarly literature on dual process accounts.

This argument applies to conflict resolution models only i.e. both Bobbie and Joanna decide to participate in all thematic rounds of the pub quiz. There is no situation where one assumes a passive role and takes a backseat. This has also been the common theme of most psychological theorizing; as Kahneman (2012, 24) writes, ‘in the story that I will tell, systems 1 and 2 are both active whenever we are awake’. There is, however, more to cognitive processes than just following a sequential, default intuition and interventionist analytic structure. While parallel competitive models have garnered attention alongside the sequential models, a hybrid set up is yet to yield empirically robust results. What has been the cornerstone of most behavioural instruments was that they have encountered scenarios in which it was deemed suitable to endorse either the parallel competitive or the default-interventionist framework of the dual process theories. However, with evidence from the cognitive psychology literature and neuroscientific findings that different sections of the brain might be activated in response to a common stimuli, there is the possibility of proposing a device that might be a hybrid; one that essentially combines the two cognitive processes, and shares features of both the parallel competitive and default interventionist dual processes accounts. The upcoming section will outline the operationalisation of nudge plus before

comparing the different behavioural instruments, and then outline testable predictions before concluding.

Operationalising nudge plus

How can nudge plus be administered? Recall the pub quiz. In their team, Bobbie and Joanna were free to adopt different conflict resolution strategies; what was essential is how they, as a team, decided to resolve such conflicts. The brain as a whole, also functions as a team when faced with decision-making choices. Put simply, the two types of brain processes are functional members of this team.

Faced with a choice and when in conflict, either Bobbie or Joanna can have a verbal duel and the emerging winner provides with the team's answer; or, they could begin with a rule-based system where one team member remains dominant and validates (or, rectifies) the response generated by the other member. The former, the parallel competitive model, suggests that a simultaneous nudge plus can be administered, where in essence, the nudge and its plus are delivered such that they prompt both type of brain processes to act together. The latter supports a sequential nudge plus mechanism; where the plus precedes or follows the nudge, letting automatic and reflective processes to act in sequence, reinforcing each other's effects. The plus can be conceived by the policy maker to be delivered before, after or as part of the classic nudge.

The order in which the nudge and the plus act are, however, context relevant to the task (policy challenge) at hand and has implications on what the plus seeks to achieve as part of its delivery. While a nudge plus promises greater, if not equivalent, autonomy relative to classic nudges, each works differently on accounts of transparency

and effectiveness. Some classes of plus, for instance, may work by making the existing nudge more salient to the receiver and thereby increasing the effectiveness of the nudge only and keep transparency unchanged; other classes may work by making the nudge construct more transparent to the receiver. Others might allow the agent to reflect on their actions completely, thereby making it even more transparent. However, the change in effectiveness might be a priori ambiguous to the policy maker in these latter instances.

A classic traffic lighting scheme, for instance, when combined and delivered with salience building labelling information can be classified to be a simultaneous nudge plus that offer greater effectiveness. While the construct of a traffic lighting nudge rests on the tenet that the agent is subtly reminded of red lights meaning ‘stop’, green lights meaning ‘go’ and amber ones meaning ‘at your own risk’, such that they make the healthier and safer lifestyle choices automatically; agents, with strong antecedent preferences or acting in weakly ecological-rational settings, might miss the visual cue, thereby, rendering the nudge ineffective. In these instances, adding an information label explicitly spelling out what the colour coding means would draw the agent’s attention, and lead to a greater uptake of the nudge while guaranteeing enhanced autonomy as the agent can selectively decide to ignore the nudge. Similarly, providing a choice to commit before or after a default setting could have different implications. A default setting is usually taken up by an agent due to the cognitive easing it comes with; for instance, choosing from a set-menu is often easier for an individual who dreads a large menu with many options. In this setting, providing the agent with the choice to commit to a healthier diet before the default menu is presented makes the uptake of the default more salient to the agent suggesting greater effectiveness. However, if the sequence of this nudge plus is reverted such that the set menu is

presented first, and just before ordering the agent is asked to commit to a healthier diet, the choice thereafter would be governed by self-reflection, and in essence should be more transparent to the agent, even though the treatment effect of the nudge plus might be less effective than its opaque nudge counterpart.

Generally speaking, the role plus plays and outcomes it aims to achieve, in turn depends on kind of reflection that the plus seeks to deliver. Reflection as a means of learning has been researched quite extensively (for a review, see Atkins and Murphy 1993). While the idea of reflection and reflective strategies is abstract at its best, two strands of literature can be held accountable for the promising leap in understanding what it means; one can be traced back to educational theorists who postulated as early as the 1980s, the role reflection takes on in reforming one's own self and behaviour, a form of conscious or deliberate learning (Tough 1979). The other strand can be traced back to the dual process theorists who have, timelessly, referred to the effortful cognitive processes as the reflective brain processes. In the account that we aim to provide here, we abide by dual process view that type 2 processes are capable of reflection and what they lead to in the conceptualisation of nudge plus is an 'experiential learning' environment, which if accepted by the receiver leads to a persistent effect of the classic nudge.

Reflection, true to its origin as seeing 'one's reflection in the mirror' has often been used to relate to self-reflection. An intellectual, for instance, is a mind that watches itself. However, contrary to the common belief, reflection could personify different meanings; for instance, one could reflect on one's own choices, beliefs, thoughts or feelings, or one could reflect on the available alternatives, or even more could reflect on the structural assumptions behind a construct. To be precise, reflection involves thinking about something. However, is reflection always conscious? Reflection could

be conscious, for instance, when one deliberately thinks of options or unconscious when triggered by the environmental cues. Even more, it could be a solitary act when one reflects on the choices, or as part of the herd; for instance, debriefing exercises help reflection (Boud 1985). However, whether self or group led, conscious or unconscious, reflection involves transformation of perspectives: old to new.

However, what does this transformation entail? Different scholars have recognised different stages in reflection that lead to the generation of new perspectives; Schon (1991), for instance, identifies three different stages in reflection: think, criticize, and acting accordingly. Mezirow (1981; 1998) accounts for a seven-stage reflective process. However, as Atkins and Murphy (1993) put it, reflection involves sharing an initial discomfort/dissonance from a certain stimuli, followed by critical analysis of one's feelings and thereafter acting in accordance to them. This critical reflection could either come as a sudden shock, for instance, a heart disorder can cause reflective processes to be activated leading to a change in lifestyle; or, it can come in transition, for instance, borderline changes in blood sugar levels can make someone conscious of their lifestyle habits and introduce small lifestyle changes. However, one must be motivated, conscientious and goal-oriented to act on the cues and reflect choices. This distinguishes reflection from just letting thoughts emerge. Considering this, reflection as the plus, can be defined as follows:

Reflection is the act of thinking and re-evaluating prior actions, choice constructs or available alternatives, triggered by conscious or unconscious experiences, that when engaged in by the agent results in the uptake of newer perspectives.

The plus, when taken up by the agent, implies self-reflection in relation to long-term preferences, for instance, when asked to commit after a default, or a deeper thinking about the choice construct of the behavioural tool at hand, or where greater salience can lead to an enhanced uptake of the nudge or a greater transparency into the decision making scenario or both. Whatever might be the reflective process, the agent learns in the environment and behaves accordingly with a new reactance. The environment, in turn, is either transparency enhancing, or effectiveness enhancing, or both. As such, nudge plus can have an effect through both the nudge and its plus, but to reinforce a persistent change, the policy maker must decide on whether the plus comes before, after or with the nudge depending on the context in hand. Most domains of nudge could include a reflective component, even the default, as the dialogue between the policy-maker and the citizen is not a one off, but occurs over time, perhaps over the whole lifespan of a citizen, such as a pensions default that occurs at the start of someone's career but is explained at a later stage.

To sum up, nudge plus mixes a classic nudge with reflective strategies. However, given the dual process accounts that we have discussed, such combinations could either be simultaneous or sequential in nature; simultaneous meaning both the nudge and plus are administered at the same time, and sequential meaning one is preceded by the other in any logical order. The operationalisation of nudge plus is supported by cognitive dual process accounts; when the nudge plus is simultaneously administered it would closely resemble the account of a parallel competitive dual processes model, while in its application as a sequential nudge plus, it would mimic the default-interventionist approach.

Comparing nudge, nudge plus, and boost

Now that we have set out an account of cognition, how do these understandings play out with the three kinds of intervention we started with, nudge, nudge plus, and boost? To begin with, here is a simple decision-making task to help understand how nudge, nudge plus, and boost might differ by their functionality.

Consider visiting a restaurant and placing an order. This entails a simple search to draw information from the underlying characteristic set, consisting of agent relative traits and environmental factors, and use simple search rules R_1 to narrow down the alternatives. The search rules, for instance, in this exemplar could be to look for a restaurant nearest to my place or one that serves a particular cuisine. The search leads to a set of available alternatives which then have underlying properties; for instance, price, quality of dish served and so on which through a series of selection rules R_2 are then narrowed down to a final choice. Let Figure 1 represent this choice mechanistic scheme. Now in this given context, the behavioural instruments might work very differently.

Nudges, for instance, will work by co-opting the biases of an individual and changing the choice construct only, indicated by path line L_1 , such that all other attributes including the set of alternatives and their properties remain unchanged. Boosts, on the contrary, work by improving the competencies of the agent. As such, the boosts will change the underlying search and selection rules, R_1 and R_2 only, keeping the choice environment and the alternatives unchanged. These changes in the rules can be at times driven by changes in the informational environment, those referred to by Herwig and Grüne-Yanoff (2017) as short-term boosts but that is not all of it. Furthermore, any regulatory policy would involve changing the set of alternatives or

its properties ideally; for instance, a ban would reduce the set of alternatives or a price change would change the properties set.

Nudge plus could work differently to all of these indicated paths. This, however, is contingent on what role the plus plays. Consider, for instance, a plus that makes the nudge more salient by making the choice construct clearer. This involves path L_1 but simultaneously draws on the environment's informational subset as well. Similarly, when the plus embodies a reflection that allows the agent to reflect on the available alternatives, the plus involves re-evaluating the set of alternatives either before, after or along with the path L_1 . Now what happens, if one self-reflects on their actions? This involves drawing on the agent-relative properties, besides affecting L_1 .

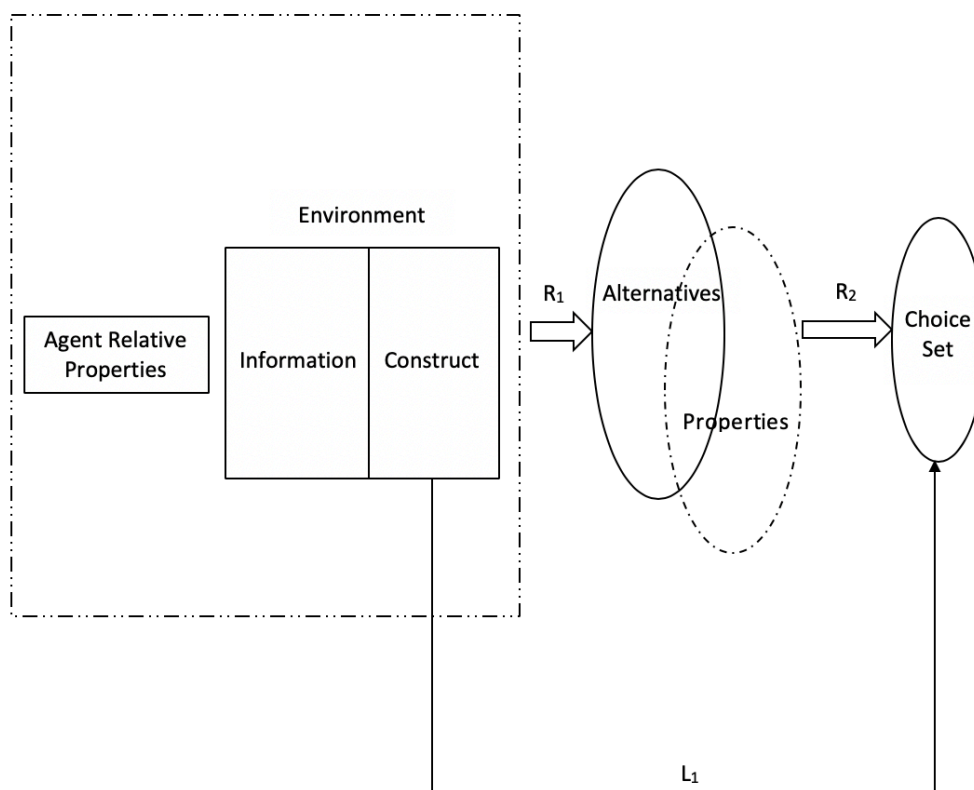


Figure 1: A mechanistic scheme³

³ This representation of a mechanistic scheme was adapted from one presented by Till Grune-Yanoff in a lecture on the 27 November 2019.

Nudge plus works by building on choice constructs of the nudges, while boosts and other instruments adopt a different causal pathway. These functional differences can, however, be explained by a range of operational parameters, to which we now turn. Table 1 does the job of comparing them and showing how they depend on different accounts of cognition. If we accept the argument of the previous section that a more sophisticated account of cognition is needed, then the case improves for nudge plus.

With respect to the psychological paradigm, although both the heuristics and biases (H&B) and the simple heuristics (SH) approach begin with heuristics as their cognitive foundation, they differ in their conceptual rationale in justifying humans' bounded rationality and their associated failures. While the H&B presumes that all biases are systematic, and that they are usually generalizable across the population in a way that a bias will arise when there is a heuristic that is adopted; SH, contrarily, denies this by acknowledging that humans follow short cuts, and they might go wrong at times, but this is not systematic. In doing so, the proponents of the SH approach believe that there remains no need to co-opt an agent's heuristics, and that their decision making can be improved by simply enlarging their adaptive toolbox. Given that classic nudges and boosts belong to these two different psychological schools of thought, it can be easily anticipated that their underlying processing (cognitive) architectures will differ even if they lead to the same behavioural outcome.

In line with our conceptualisation of the nudge plus, there is sufficient reason to believe that it is closely aligned to the H&B approach. The justification for this psychological theorizing of nudge plus in turn leads us on to this discussion's second tenet of comparison: the cognitive architecture. Nudge plus remains as extension of the classic nudges and build on them by adding self-reflection strategies. In doing so, the nudge plus construct acknowledges the validity of the dual process accounts as because,

by conceptualisation, the nudge component functions by taking advantage of an agent's volitional biases that co-occur with her heuristically based decisions facilitated by the cue-based autonomous (intuitive) responses; while the plus component induces her reflective processes to work which can thereafter reinforce the behavioural change incorporated by the nudge.

It is important to stress that nudge plus restores the autonomy of the agent, which addresses the risk of the reversing the behavioural change and improves the transparency of instituting such changes, addressing in part some of the ethical objections to nudges. The agent can decide whether such changes are compatible to other decisions and preferences which is more likely to lead to persistent behavioural changes. Even though the classic nudge, through repeated applications, can also sustain behaviour over time, nudge plus may achieve such persistent behavioural changes even as a one-off application. Self-realisation of the nudge's objective by the agents, may embed transformed behaviour in a new habitual pattern, sustaining it even after the withdrawal of the nudge plus.

The final theme of comparison between the behavioural change seeking instruments: the motivation and competence of the decision-making agent. The problem identified with the incorrectly designed commitment device can also be extended to a well-designed policy whether a plus or boost. This has been well documented by Hertwig and Grune-Yanoff (2017) as the policy-theory coherence dilemma. If agents lack motivation, a boost and plus will under achieve their targets. In the worst possible scenario failing reflection on lacking motivation, nudge plus would end up delivering the same effect as a classic nudge (just the way a boost or think fails). Set out in this way, this creates the testable implication that an element of reflection as part of the delivery of the nudge will improve outcomes compared to the

classic nudge. We cannot adjudicate from theory between the effectiveness of nudge plus and boost or whether the plus or boost is better delivered before the intervention or during it. We can hypothesise that a nudge plus will beat a boost and the classic nudge when outcomes are examined over the long term and in successive interventions because of the self-knowledge that nudge plus entails. Nudge plus beating nudge, and over the long-term beating boost, are two direct testable implications we draw. Furthermore, the plus might also lead to promoting behavioural spill overs in other domains as agents learn to reflect on life-choices in general and is the third testable prediction.

Dimension	Classic nudge	Nudge Plus	Boost
<i>Psychological Paradigm</i>	Heuristics and Biases	Heuristics and Biases	Simple Heuristics
<i>Cognitive Architecture</i>	Dual Process Theory	Dual Process Theory	Malleable cognitive structure
<i>Reversibility</i>	Reversible	Persistent effects	Persistent effects
<i>Opacity</i>	Usually opaque	Transparent with the plus element	Completely transparent
<i>Autonomy</i>	No autonomy	Autonomy comes with the plus element	Complete autonomy
<i>Causal Pathway</i>	Behaviour	Behaviour	Competency
<i>Bias Awareness and Control</i>	No Awareness and Control	Control given with the plus element	Agent is aware and in control of the biases
<i>Social Planner's information about end goals and benevolence</i>	Social planner is assumed to be benevolent and aware of end goals	Social planner must be aware of end goals but decision is left to the agent. Can be rent seeking	No need for social planner to be aware of the goals. Social planner can be rent-seeking
<i>Cognitive error of Social planner</i>	Must not be error prone	Can be error prone	Can be error prone
<i>Motivation and competence of decision maker</i>	Not required. The decision maker is a cognitive cripple	The decision maker must be goal-oriented, conscientious and motivated to act.	The decision maker must be competent and motivated to act.
<i>Examples</i>	Opt-out Defaults, Traffic Lighting Scheme, Commitment Devices, et cetera	Nudges like Defaults or Traffic lighting schemes combined with pluses like Active Decision Mechanisms like contracts or personal budgeting schemes or salience building information or social feedback schemes that can led to reflection.	Educative Nudges, Fast and Frugal Trees, Quick Rules, Numeracy enhancing skills, Implementation Intentions

Note: Columns 1 and 3 as adopted from Hertwig and Grüne-Yanoff (Table 1, p16, 2016; Table 1, p2, 2017)

Table 1: Conceptual Underpinnings of different behavioural policies

Conclusion

The aim of this paper is to develop a modified account of behaviour change—nudge plus—based on the idea that encouraging an element of reflection as part the delivery of the nudge would enhance outcomes. Nudge plus gives an opportunity for citizens to own the process; and thereby they commit and invest in it. A key further aim in this paper is to distinguish between strategies of behaviour change, and the most powerful way of doing it is with a plausible account of dual processes. The review of dual systems shows that the pure dependence on dual systems implied by classic nudge is not sustainable at least in all domains; there is a better case for a more plastic account of the two systems implied by boosts. The idea that the two systems are connected is support for nudge plus as it is based on a plausible account of cognition.

Having set out the conceptual foundations, we have elaborated the potential for nudge plus as a form of behavioural public policy. We have set out testable claims of the differences between the three kinds of intervention. There remains of course more conceptual work to do, such as to examine the welfare implications of increasing autonomy with nudge plus, and to find out the extent to which individuals are still being manipulated even if they are encouraged to reflect as it might be clear what direction the reflection is supposed to lead. Also nudge plus and boosts have been presented as competitors as forms of intervention, but they may be complementary and work more strongly in tandem with each other. Even with these further theoretical challenges, the way forward at the moment is for more tests of nudge plus to take place.

References

- Anonimo, Jonathan St B. T. Evans, Stephen E. Newstead, and Ruth M. J. Byrne. 1993. *Human Reasoning: The Psychology of Deduction*. Psychology Press.
- Atkins, Sue, and Kathy Murphy. 1993. 'Reflection: A Review of the Literature'. *Journal of Advanced Nursing* 18 (8): 1188–92. <https://doi.org/10.1046/j.1365-2648.1993.18081188.x>.
- Baldwin, Robert. 2014. 'From Regulation to Behaviour Change: Giving Nudge the Third Degree'. *The Modern Law Review* 77 (6): 831–57. <https://doi.org/10.1111/1468-2230.12094>.
- Bargh, John A. 1994. 'The Four Horsemen of Automaticity: Awareness, Intention, Efficiency, and Control in Social Cognition'. In *Handbook of Social Cognition*, edited by R. Wyer and T. Srull. Lawrence Erlbaum.
- Blattman, Christopher, Julian C. Jamison, and Margaret Sheridan. 2017. 'Reducing Crime and Violence: Experimental Evidence from Cognitive Behavioral Therapy in Liberia'. *American Economic Review* 107 (4): 1165–1206. <https://doi.org/10.1257/aer.20150503>.
- Boud, David, ed. 1985. *Reflection: Turning Experience into Learning*. London: Kogan Page [u.a.].
- Bovens, Luc. 2008. 'The Ethics of Nudge'. In *Preference Change: Approaches from Philosophy, Economics and Psychology*, edited by Mats J. Hansson and Till Grüne-Yanoff, 207–20. Berlin: Springer, Theory and Decision Library A.
- Bradt, Jacob. 2019. 'Comparing the Effects of Behaviorally Informed Interventions on Flood Insurance Demand: An Experimental Analysis of “Boosts” and “Nudges”'. *Behavioural Public Policy*, October, 1–31. <https://doi.org/10.1017/bpp.2019.31>.
- De Neys, Wim. 2012. 'Bias and Conflict: A Case for Logical Intuitions'. *Perspectives on Psychological Science* 7 (1): 28–38.
- . 2014. 'Conflict Detection, Dual Processes, and Logical Intuitions: Some Clarifications'. *Thinking & Reasoning* 20 (2): 169–187.
- De Neys, Wim, and Gordon Pennycook. 2019. 'Logic, Fast and Slow: Advances in Dual-Process Theorizing'. *Current Directions in Psychological Science* 28 (5): 503–509.
- Drexler, Alejandro, Greg Fischer, and Antoinette Schoar. 2014. 'Keeping It Simple: Financial Literacy and Rules of Thumb'. *American Economic Journal: Applied Economics* 6 (2): 1–31. <https://doi.org/10.1257/app.6.2.1>.
- Epstein, S. 1994. 'Integration of the Cognitive and the Psychodynamic Unconscious'. *American Psychologist* 49 (8): 409–24.
- Evans, Jonathan. 2007. 'On the Resolution of Conflict in Dual Process Theories of Reasoning'. *Thinking & Reasoning* 13 (4): 321–39. <https://doi.org/10.1080/13546780601008825>.
- . 2010. *Thinking Twice: Two Minds in One Brain*. Oxford, New York: Oxford University Press.
- Evans, Jonathan, and Keith E. Stanovich. 2013. 'Dual-Process Theories of Higher Cognition Advancing the Debate'. *Perspectives on Psychological Science* 8 (3): 223–241. <https://doi.org/10.1177/1745691612460685>.
- Frankish, Keith. 2010. 'Dual-Process and Dual-System Theories of Reasoning'. *Philosophy Compass* 5 (10): 914–26. <https://doi.org/10.1111/j.1747-9991.2010.00330.x>.

- Gawronski, Bertram, and Laura A. Creighton. 2013. 'Dual Process Theories'. *The Oxford Handbook of Social Cognition*, August.
<https://doi.org/10.1093/oxfordhb/9780199730018.013.0014>.
- Gronchi, Giorgio, and Fabio Giovannelli. 2018. 'Dual Process Theory of Thought and Default Mode Network: A Possible Neural Foundation of Fast Thinking'. *Frontiers in Psychology* 9. <https://doi.org/10.3389/fpsyg.2018.01237>.
- Grüne-Yanoff, Till, and Ralph Hertwig. 2016. 'Nudge versus Boost: How Coherent Are Policy and Theory?' *Minds and Machines: Journal for Artificial Intelligence, Philosophy and Cognitive Science* 26 (1–2): 149–83.
<https://doi.org/10.1007/s11023-015-9367-9>.
- Heller, Sara B., Anuj K. Shah, Jonathan Guryan, Jens Ludwig, Sendhil Mullainathan, and Harold A. Pollack. 2017. 'Thinking, Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago'. *The Quarterly Journal of Economics* 132 (1): 1–54. <https://doi.org/10.1093/qje/qjw033>.
- Hertwig, Ralph, and Till Grüne-Yanoff. 2017. 'Nudging and Boosting: Steering or Empowering Good Decisions'. *Perspectives on Psychological Science* 12 (6): 973–86. <https://doi.org/10.1177/1745691617702496>.
- Jacoby, Larry L. 1991. 'A Process Dissociation Framework: Separating Automatic From Intentional Uses of Memory'. *Journal of Memory and Language* 30: 513–41.
- John, Peter. 2018. *How Far to Nudge?: Assessing Behavioural Public Policy*. Cheltenham, UK ; Northampton, MA: Edward Elgar Publishing Ltd.
- John, Peter, Sarah Cotterill, Alice Moseley, Liz Richardson, Graham Smith, Gerry Stoker, and Corinne Wales. 2019. *Nudge, Nudge, Think, Think: Experimenting with Ways to Change Citizen Behaviour*,. S.l.: Manchester University Press.
- John, Peter, Sarah Cotterill, Liz Richardson, Alice Moseley, Graham Smith, Gerry Stoker, and Corinne Wales. 2013. *Nudge, Nudge, Think, Think: Experimenting with Ways to Change Civic Behaviour*. A&C Black.
- John, Peter, and Gerry Stoker. 2019. 'Rethinking the Role of Experts and Expertise in Behavioural Public Policy'. *Policy & Politics* 47 (April).
<https://doi.org/10.1332/030557319X15526371698257>.
- Kahneman, Daniel. 2012. *Thinking, Fast and Slow*. London: Penguin.
- Kahneman, Daniel, and Shane Frederick. 2002. 'Representativeness Revisited: Attribute Substitution in Intuitive Judgment'. In *Heuristics and Biases: The Psychology of Intuitive Judgment*, 49–81. New York, NY, US: Cambridge University Press. <https://doi.org/10.1017/CBO9780511808098.004>.
- Kahneman, Daniel, Stewart Paul Slovic, Paul Slovic, and Amos Tversky. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge University Press.
- Klaczynski, Paul A. 2000. 'Motivated Scientific Reasoning Biases, Epistemological Beliefs, and Theory Polarization: A Two-Process Approach to Adolescent Cognition'. *Child Development* 71 (5): 1347–1366.
- Kruglanski, Arie W., and Gerd Gigerenzer. 2011. 'Intuitive and Deliberate Judgments Are Based on Common Principles'. *Psychological Review* 118 (1): 97–109.
<https://doi.org/10.1037/a0020762>.
- Lieberman, Matthew D. 2003. 'Reflexive and Reflective Judgment Processes: A Social Cognitive Neuroscience Approach'. In *Social Judgments: Implicit and Explicit Processes*, 44–67. New York, NY, US: Cambridge University Press.
- Lindsay, D. Stephen, and Larry L. Jacoby. 1994. 'Stroop Process Dissociations: The Relationship between Facilitation and Interference'. *Journal of Experimental*

- Psychology: Human Perception and Performance* 20 (2): 219–34.
<https://doi.org/10.1037/0096-1523.20.2.219>.
- Loewenstein, Goerge, Cindy Bryce, David Hagmann, and Sachin Rajpa. 2015. 'Warning: You Are About to Be Nudged'. *Behavioral Science and Policy* 1 (1): 35–42.
- Marteau, Theresa M., David Ogilvie, Martin Roland, Marc Suhrcke, and Michael P. Kelly. 2011. 'Judging Nudging: Can Nudging Improve Population Health?' *BMJ (Clinical Research Ed.)* 342 (January): d228.
- Mezirow, Jack. 1981. 'A Critical Theory of Adult Learning and Education': *Adult Education*. <https://doi.org/10.1177/074171368103200101>.
- . 1998. 'On Critical Reflection': *Adult Education Quarterly*. <https://doi.org/10.1177/074171369804800305>.
- Mols, Frank, S. Alexander Haslam, Jolanda Jetten, and Niklas K. Steffens. 2015. 'Why a Nudge Is Not Enough: A Social Identity Critique of Governance by Stealth'. *European Journal of Political Research* 54 (1): 81–98.
<https://doi.org/10.1111/1475-6765.12073>.
- Mühlböck, Monika, Fabian Kalleitner, Nadia Steibe, and Bernhard Kittel. 2019. 'Information, Reflection, and Successful Job Search: A Nudging Experiment'.
- Oliver, Adam. 2017. *The Origins of Behavioural Public Policy*. <https://doi.org/10.1017/9781108225120>.
- Osman, Magda. 2004. 'An Evaluation of Dual-Process Theories of Reasoning'. *Psychonomic Bulletin & Review* 11 (6): 988–1010.
<https://doi.org/10.3758/BF03196730>.
- Paap, Kenneth R., Ronald W. Noel, and Linda S. Johansen. 1992. 'Chapter 15 Dual-Route Models of Print to Sound: Red Herrings and Real Horses'. In *Advances in Psychology*, edited by Ram Frost and Leonard Katz, 94:293–318. Orthography, Phonology, Morphology, and Meaning. North-Holland.
[https://doi.org/10.1016/S0166-4115\(08\)62800-9](https://doi.org/10.1016/S0166-4115(08)62800-9).
- Pollan, Michael. 2010. *Food Rules: An Eater's Manual*. Camberwell, Vic.: Penguin.
- Schon, Donald A. 1991. *The Reflective Practitioner: How Professionals Think in Action*. Ashgate.
- Sherman, Jeffrey W. 2006. 'TARGET ARTICLE: On Building a Better Process Model: It's Not Only How Many, but Which Ones and By Which Means?' *Psychological Inquiry* 17 (3): 173–84.
https://doi.org/10.1207/s15327965pli1703_3.
- Stanovich, Keith E., and Richard F. West. 2000. 'Individual Differences in Reasoning: Implications for the Rationality Debate?' *Behavioral and Brain Sciences* 23 (5): 645–65. <https://doi.org/10.1017/S0140525X00003435>.
- Stutzer, Alois, Lorenz Goette, and Michael Zehnder. 2011. 'Active Decisions and Prosocial Behaviour: A Field Experiment on Blood Donation'. *The Economic Journal* 121 (556): F476–93. <https://doi.org/10.1111/j.1468-0297.2011.02477.x>.
- Sunstein, Cass R. 2016a. 'People Prefer System 2 Nudges (Kind Of)'. SSRN Scholarly Paper ID 2731868. Rochester, NY: Social Science Research Network. <https://papers.ssrn.com/abstract=2731868>.
- . 2016b. *The Ethics of Influence: Government in the Age of Behavioral Science*. Cambridge University Press.
- Sunstein, Cass R., and Richard H. Thaler. 2003. 'Libertarian Paternalism Is Not an Oxymoron'. *The University of Chicago Law Review* 70 (4): 1159–1202.
<https://doi.org/10.2307/1600573>.

- Thaler, Richard H., and Cass R. Sunstein. 2009. *Nudge: Improving Decisions About Health, Wealth and Happiness*. London: Penguin.
- Tough, Allen M. 1979. *The Adult's Learning Projects: A Fresh Approach to Theory and Practice in Adult Learning*. Ontario Institute for Studies in education.
- Zimmerman, Frederick J. 2016. *From Choice Architecture to Policy Infrastructure: Multilevel Theory and the Political Economy of Health Behaviors*. Oxford University Press.