

University of Groningen

The ethics of nudging

Schmidt, Andreas T.; Engelen, Bart

Published in:
Philosophy Compass

DOI:
[10.1111/phc3.12658](https://doi.org/10.1111/phc3.12658)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2020

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Schmidt, A. T., & Engelen, B. (2020). The ethics of nudging: An overview. *Philosophy Compass*, 15(4), [e12658]. <https://doi.org/10.1111/phc3.12658>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

ARTICLE

The ethics of nudging: An overview

Andreas T. Schmidt¹  | Bart Engelen²¹University of Groningen²Tilburg University**Correspondence**

Andreas T. Schmidt, Faculty of Philosophy,
University of Groningen, Oude
Boteringestraat 52, 9712 GL Groningen, The
Netherlands.
Email: a.t.schmidt@rug.nl

Abstract

So-called nudge policies utilize insights from behavioral science to achieve policy outcomes. Nudge policies try to improve people's decisions by changing the ways options are presented to them, rather than changing the options themselves or incentivizing or coercing people. Nudging has been met with great enthusiasm but also fierce criticism. This paper provides an overview of the debate on the ethics of nudging to date. After outlining arguments in favor of nudging, we first discuss different objections that all revolve around the worry that nudging vitiates personal autonomy. We split up this worry into different dimensions of autonomy, such as freedom of choice, volitional autonomy, rational agency, and freedom as nondomination. We next discuss worries that nudging is manipulative, violates human dignity, and prevents more important structural reform. Throughout, we will present responses that proponents of nudging can muster. On the whole, we conclude that the objections fail to establish that the nudge program as a whole should be rejected. At the same time, they give us important guidance when moving towards an ethical assessment of nudges on a case-by-case basis. Towards the end, we provide some possible ways forward in debates around the ethics of nudging.

1 | INTRODUCTION

Since the publication of Richard Thaler's and Cass Sunstein's much celebrated book *Nudge*, many policymakers, politicians, and academics have enthusiastically embraced the use of behavioral science in public policy (Thaler &

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2020 The Authors. *Philosophy Compass* published by John Wiley & Sons Ltd

Sunstein, 2008). Nudging holds a powerful promise: Rather than changing people's options or economic incentives, we can improve people's decisions by changing how options are presented to them. Since the first articles on the ethics of nudging appeared around a decade ago, a rich ethical debate has emerged too, with sophisticated arguments for and against nudging (Bovens, 2009; Hausman & Welch, 2010). In this overview article, we first describe what nudging is (Section 2) to then discuss the ethical considerations for and against it (Sections 3 and 4). In the end, we conclude and propose some ways forward for future enquiries into the ethics of nudging (Section 5).

2 | WHAT IS NUDGING?

Thaler and Sunstein (2008, p. 6) define a nudge as “any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives.” Nudges then refer to the deliberate changes to and designs of people's choice environments—the ways in which options are presented or framed—in attempts to predictably steer those people in specific directions. Nudging has been received with much enthusiasm in public policy.¹ A common application are defaults, such as an opt-out system for organ donation to increase the proportion of potential donors by relying on people's predictable tendencies to stick to the status quo (Rithalia, McDaid, Suekarran, Myers, & Sowden, 2009). Or take the smart design of cafeterias: Putting healthy food options such as apples at eye level boosts apple consumption because salient items predictably capture the attention of customers (Arno & Thomas, 2016). Other nudges work through visual designs. Horrible pictures on cigarette packs, for example, reduce cigarette consumption through emotional responses (Fong, Hammond, & Hitchman, 2009). Smart road markings can reduce traffic accidents in multiple ways (Carlson, Park, & Andersen, 2009).

Note three things about our working definition of nudging.

First, on a common characterization, nudges typically work by triggering or tapping into cognitive heuristics and other so-called “System 1” mechanisms. These psychological mechanisms are typically faster and less conscious than the slower and more conscious reflective capacities of “System 2,” which are the psychological mechanisms active when people deliberate, justify their beliefs and actions, and weigh off reasons (Kahneman, 2011). While such a “dual processing” theory is controversial and arguably provides an oversimplified picture of people's mental capacities, it is still useful as a point of departure and has broadened our thinking about how to influence people without merely informing or persuading them (Thaler & Sunstein, 2008, p. 19). Alternative theories, such as the “fast and frugal heuristics” research program, also refer to the important role cognitive heuristics play in influencing beliefs and actions (see, e.g., Gigerenzer and Brighton, 2009). In our view, merely providing reasons or information would not qualify as a nudge for two reasons (Sunstein, 2015b). First, most people distinguish nudges as novel policy tools different from more traditional tools such as providing information and incentivizing and coercing people. Second, merely providing information does not raise the kind of ethical worries that critics of nudging have voiced and which we discuss below.

Second, our working definition does not specify the directions people are nudged into. Ethical discussions often focus narrowly on paternalist nudges which aim to make nudgees better off. But many nudges are not paternalist but instead benefit the environment (e.g., painted footsteps towards bins for recycling), benefit others (e.g., opt out systems for organ donation), or reduce the risk posed to others (e.g., smart road markings for safer driving). Accordingly, we do not limit our discussion to issues around paternalism.

Finally, we here do not discuss behavioral policies that use economic incentives. Think of paying mothers for quitting cigarettes during pregnancy. Such interventions do not qualify as nudges, narrowly conceived, because they do change the incentives associated with options. At the same time, they resemble nudges as they use behavioral insights about people's decision-making propensities that deviate from the traditional *homo economicus* perspective. Accordingly, even though we do not discuss such policies here, the ethical issues we explore apply, *mutatis mutandis*, to such behavioral policies too.

3 | ARGUMENTS FOR NUDGING

Several straightforward arguments speak for nudging.

First, the nudge program promises cost-effective policies to promote beneficial policy outcomes. Insights in behavioral science afford policymakers with new tools to, among other things, promote better health behavior, pro-environmental behavior, and higher tax compliance. Moreover, nudges are comparatively cheap and easy to implement.

Of course, many nudge interventions fail to be effective. But such findings do not invalidate the nudge approach as a whole. As long as enough nudge interventions work—as seems to be the case—and we can separate effective from ineffective nudge policies, the nudge approach can still produce valuable policies (Arno & Thomas, 2016; Hummel & Maedche, 2019). Nudges typically lend themselves nicely to an evidence-based approach to policymaking and can often even be tested through randomized control trials. Moreover, proponents of nudging, such as Sunstein, also tend to endorse a stronger evidence orientation in public policy (Sunstein, 2019). That we regularly find ineffective nudges arguably indicates that the approach does reasonably well in separating effective from ineffective policies.

Second, nudges typically respect freedom of choice, as they do not change or remove options nor significantly change economic incentives. Freedom of choice can be valuable as it is often necessary to respect autonomy (see Section 4.1). But freedom of choice might also matter politically, as nudging might sidestep some partisan and ideological disagreements that so often beset politics.

Third, some evidence suggests that, compared with traditional interventions like taxation and fines, citizens are more likely to accept nudges. On the whole, people typically tend to welcome nudges, at least insofar as they promote goals they endorse and are implemented by parties with whom they can identify (Hagman, Andersson, Västfjäll, & Tinghög, 2015; Petrescu, Hollands, Couturier, Ng, & Marteau, 2016; Reisch & Sunstein, 2016; Reisch, Sunstein, & Gwozdz, 2017; Sunstein, 2016, Chapter 6).

A final argument is that choice architecture seems inevitable (Sunstein, 2015a, pp. 420–422). Frequently, there is no neutral way of framing options and decisions will always be influenced one way or another. For example, defaults or physical choice architectures are often inevitable. Refraining from deliberating designing defaults and choice environments would then be a choice to ignore how those might lead people towards suboptimal choices.

However, the inevitability argument alone is unlikely to win over all critics (Vallier, 2016). Distinguish between choice architecture that happens somewhat unintentionally or even randomly and choice architecture that was deliberately designed (Hausman & Welch, 2010, p. 133). Policy nudges feature intentional choice architecture and as such might raise ethical worries that do not apply to random influences. Some of the worries we discuss below, for example, surrounding manipulation and power abuse, only apply to intentional nudges. Proponents of nudging should, we think, acknowledge that intentional nudges raise ethical challenges. At the same time, as will become clear below, much of the choice architecture placed by nongovernmental “architects,” such as private companies and marketeers, is intentional too. Accordingly, the choice we often face is not between “intentional” and “unintentional” choice architecture but between different kinds of intentional choice architecture.

In sum, nudging offers the promise of cost-effective, evidence-driven, freedom-preserving, nonpartisan, and comparatively popular policies. Moreover, as choice architecture often influences us one way or another, why not utilize it to make us better off?

4 | ARGUMENTS AGAINST NUDGING

Amidst great enthusiasm, nudging has also received much criticism. Critical voices hold that nudging can clash with central moral values, like liberty, autonomy, respect, and dignity. We here present oft-heard objections to nudging and outline comebacks in defense of nudging. Note that the objections we discuss mostly argue that there is something problematic about the *nudge approach* as a whole rather than just about isolated nudge instances.

4.1 | Autonomy

Most ethical worries revolve around autonomy. “Autonomy” is a notoriously rich notion, so we here separate it along four dimensions to better tease out different objections.

- i. Freedom of choice
- ii. Psychological autonomy as volitional autonomy
- iii. Psychological autonomy as (instrumental) rational agency
- iv. Absence of domination

We now discuss different objections to nudging by going through the different dimensions.

- i. Freedom of choice

The first autonomy dimension is about the external options individuals have. Freedom is here understood as an “opportunity-concept” rather than an “exercise-concept” (Taylor, 1979). Some opponents argue that nudges can undermine freedom of choice and hence are not as “liberty-preserving” or “easily resistible” as proponents would have us believe (Grüne-Yanoff, 2012; Rebonato, 2014). But we think this worry misses the conceptual distinction between *external options*—which typically remain unaffected by nudging—and the decisions individuals end up taking (how they “exercise” their freedom). Accordingly, we think such worries are better spelled out through the other dimensions of autonomy.

- ii. Volitional autonomy

Some worries relate to *volitional autonomy*, the idea that one's actions should reflect the preferences, desires, or ends that are truly one's own. Harry Frankfurt's famous “higher-order desire” account is a conception of volitional autonomy, alongside many alternative accounts (Christman, 1991; Frankfurt, 1971; Garnett, 2014; Oshana, 2006).

The worry now is that, when we are nudged, we are no longer the “authors” of our choices: They are not really our own anymore in that they do not reflect our own autonomous desires. Nudgers pull our strings and employ tricks to get us to do what *they* want (Hausman & Welch, 2010, p. 128). When subjected to nudges, we may be influenced so that our resulting desires and actions are no longer genuinely our own.

Here are three salient responses advanced in the literature.

First, proponents of nudging could urge us to consider the counterfactual in which the nudge is removed. Removing the nudge often does not make people any more volitionally autonomous (Nys & Engelen, 2017). If we change the opt-out back to opt-in or put the apples back where they were, both the heuristics at play and many irrelevant aspects in people's choice architecture would still influence people's decisions. As mentioned above, intentional nudging can often be avoided, but choice architecture cannot. We can refrain from deliberate nudging and leave people subject to randomly designed choice architectures, for example, but it is not clear how that would better respect let alone further people's volitional autonomy.

Second, some proponents advance an even more ambitious defense: Rather than diminish volitional autonomy, nudges sometimes promote it, when they help a nudgee “choose according to her own conception of the good in situations where she would have previously done otherwise” (Mills, 2015, p. 503). A popular idea with proponents of nudging is “means paternalism.” Nudging should not be used to impose what ends people should pursue—which would be “ends paternalism”—but instead improve the means to help people achieve the ends they themselves want to pursue (Sunstein, 2016, pp. 54–56). If successful, so the argument goes, nudging can strengthen rather than undermine volitional autonomy (Engelen & Nys, 2019). (A worry is of course that nudgers might struggle to reliably identify people's ends, something we take up below in Section 4.3.)

Finally, in diverse societies, people hold very different ends and conceptions of the good. Accordingly, proponents of nudging typically hold that nudge techniques should not exert an influence so strong they are no longer “easily resistible” (Saghai, 2013, p. 489) or “easy and cheap” to avoid (Thaler & Sunstein, 2008, p. 6). Defaults, for example, in principle leave freedom of choice intact but can still have considerable impact, especially when people lack information and choice environments are confusing (Willis, 2013). Under such conditions, these techniques, proponents could argue, should then not count as nudges, as being a nudge *implies* easy resistibility. This conceptual response simply defines nudges in a way that excludes such techniques (Heilmann, 2014; Saghai, 2013). But, of course, one could simply argue that, whatever the correct definition of “nudge,” we have reason to ensure nudge interventions are easily resistible. That said, most nudging techniques are quite resistible and impact only a small percentage of the population. For a cafeteria nudge, for example, a 10% increase in apple consumption would count as a success, even though 90% will resist or ignore the nudge (see, e.g., Arno & Thomas, 2016). In a sense, “effectiveness” turns out to be a double-edged sword. Arguments for nudging require it to be effective, but make it too effective and you might end up worrying about resistibility. Yet, given the typical effect sizes of nudge policies, such worries about resistibility are unlikely to apply to most nudges.

iii. Rational agency

A slightly different worry about psychological autonomy affects the second dimension: rational agency. Rational agency is not about what ends we pursue but about how we make decisions. Some critics argue that nudges fail to respect or even positively undermine rationality, as nudges work through irrational or at least arational processes (Bovens, 2009; Conly, 2013, p. 30; Grüne-Yanoff, 2012; MacKay & Robinson, 2016). As Jeremy Waldron puts it: “I wish... I could be made a better chooser rather than having someone on high take advantage (even for my own benefit) of my current thoughtlessness and my shabby intuitions.” (Waldron, 2014) Even if nudges respect people's freedom and promote their goals and well-being, that nudgers tap into our irrational or arational heuristics and biases means they fail to treat us like rational human beings and thereby condescend and infantilize us. Instead, critics hold, policymakers should either leave us alone and let us make our own decisions (and mistakes) or pursue policies—such as educating, persuading, and informing—that treat us as rational.

A related concern here is voiced by Frank Furedi (2011), who worries that nudging can “deprive people of the capacity for making wrong choices” and erodes their responsibility for their own choices (see Waldron, 2014) for a similar objection). If responsibility for making choices is moved away from individuals, their practical judgement and decision-making capacities cannot develop, which in turn undermines their moral independence.

Here are four possible responses.

First, one can stress again that arational mechanisms are somewhat inevitable. Waldron's wish to be made a better chooser is to some extent naïve, as he assumes that he can somehow overcome or avoid the impact of arational influences. As with volitional autonomy, removing nudges will not magically restore people's rationality (Engelen, 2019b). If even subtle cues in body language influence people's behavior, can one ever avoid nudging others? “Well yes,” opponents might say, as one can avoid the *intentional* exploitation of people's foibles, something we come back to below.

Second, a proponent of nudging could also admit that nudging is not always the best policy solution. If strategies that respect or promote people's rationality, like informing or rationally persuading them, actually work, we have a *prima facie* reason to prefer those. In some cases, we can even try to “boost” their cognitive capacities instead of merely taking them as given (Barton & Grüne-Yanoff, 2015; Hertwig & Grüne-Yanoff, 2017). While opponents are more optimistic in this respect, nudge proponents typically stress scenarios where external nudges do better. If “dangerous curve ahead” signs do not prevent people from dying in car crashes, then painting white lines or visual illusions on the road is preferable. According to nudge proponents, we should treat people as they are. If people are less rational than they think or wish they were be, then respecting them as agents does not imply treating them as perfectly rational choosers (Engelen, 2019b; Schmidt, 2019).

Third, proponents of nudging could respond that critics mischaracterize how at least how some nudges work. A frequent characterization of nudges holds that they work exclusively through shallow cognitive processes and fast and automatic decision-making (what some label “System 1”) and do so opaquely. However, many nudges (at least partially) rely on deliberate and slow decision-making processes too (what some label “System 2”). Other nudges rely on both fast and frugal and deliberate and reflective processes. Hansen and Jespersen distinguish between “type 1 nudges” that mostly rely on “arational” heuristics and biases and “type 2 nudges” that appeal to both more automatic and faster and to more deliberative and slower capacities (Hansen & Jespersen, 2013; John, 2018, pp. 128–133; Schmidt, 2019). Examples of the latter are alarm sounds for seatbelts, traffic light labels for nutrients on packaging, and colorful footsteps towards recycling bins. Such nudges are transparent and direct one’s attention to relevant features in one’s choice environment. Arguably, most nudges involve at least some reflective components as well.

Finally, proponents could also put forth a more ambitious response and take issue with the notion of rationality that typically underlies the rationality objection (Schmidt, 2019). The rationality objection often draws on a view of rational agency (loosely) based on rational choice theory: Rational agents have preference orders over options, deliberately attend to relevant features of options that underlie their preference orders, and take relevant probabilities into account. For example, when choosing a retirement plan, a rational chooser will focus on the financial details of different plans. On this view, rationality is primarily located in the deliberative System 2 rather than the fast and automatic System 1. Nudges, however, seem to prioritize System 1 over System 2 and direct people towards irrelevant aspects of options rather than those that make them choiceworthy.

Now, this objection might be less forceful, if we take a different, ecological view of rationality: Different heuristics and other fast and frugal decision-making procedures can be entirely rational relative to the person who employs them and the environment in which they are used (Gigerenzer & Selten, 2002; Morton, 2011; Schmidt, 2019; Todd & Gigerenzer, 2012; Todd, Gigerenzer, and ABC Research Group, 2012). In real-life contexts, good outcomes are often better achieved, particularly under cognitive and temporal constraints, through fast and frugal heuristics instead of slow deliberation and utility maximization. Rationality, on this view, is about a good match between decision-making procedure, person, and environment. For example, in an environment in which I can rely on someone setting a default that furthers my interest, it can indeed be entirely rational to rely on a heuristic to stick with defaults. The argument then is that nudging can be compatible with rationality, if it helps achieve a better match between choice environment, agent, and decision-making procedure (Schmidt, 2019).

This argument becomes stronger if you consider how many of our environments are shaped by private companies that use choice architecture ill-matched to our decision-making procedures. Food companies, for example, systematically trigger and exploit our psychological dispositions so that we overeat and overspend. Correcting such environments through public policy nudging, or so the argument goes, could counter this and achieve a better match between our decision-making propensities and our choice environments. Nudging would thereby strengthen rather than undermine rational agency. Of course, this defense does not give policymakers a *carte blanche*: Policies other than nudging, for example, educational interventions, should often be preferred when they better improve people’s rational agency (Schmidt, 2019, pp. 535–542).

iv. Absence of domination

The final autonomy dimension is less about psychology but more about the power relationships we find ourselves in. Some critics worry that nudging can become a tool in the hands of the government to exercise problematic control over citizens and their lives. Some theorists invoke a Foucauldian framework and argue that modern governments expand their beyond traditional forms of coercive control over individuals and increasingly rely on “governmentality,” which includes the government’s power to discipline and manage individuals through their own seemingly free choices (Jones, Pykett, & Whitehead, 2011; Leggett, 2014). Other theorists rely on neo-republican theories of freedom: Nudging can furnish governments with the power to exercise dominating control over

individuals (Grüne-Yanoff, 2012; Hausman & Welch, 2010; Jones et al., 2011). Agent A's power over B is dominating, if and only if A's power is not suitably controlled by B (or by someone acting on B's behalf; Lovett, 2010; Pettit, 1997, 2014; Schmidt, 2018). The worry here is that nudging facilitates such domination, as it escapes adequate individual and democratic control, because it lacks transparency.

In response, however, one might hold that it is not obvious why nudge policies should be less transparent than other public policies and regulations. Compared with other complex policies—like taxation and regulation—nudges are comparatively easy to understand and spot. And if nudge policies are indeed amenable to being made transparent, they might be just as amenable to individual and democratic control (Ivanković & Engelen, 2019; Schmidt, 2017). But note that the claim is merely that nudging is *amenable* to democratic control, not that all current nudge practices fully meet democratic standards. For example, some worry that the part-privatization of the British nudge organization—the Behavioural Insights Unit—has led to a gap in democratic accountability (Niker, 2014; Lepenies & Matecka, 2019, pp. 9–10).

Proponents can again turn the argument around and even advance the stronger claim that public policy nudging, if done democratically, can reduce rather than increase problematic uncontrolled power. In some situations, the alternative is to leave private nudges in place. When supermarkets, fast food chains, and so on shape our food environments, they do use intentional choice architecture to nudge us towards purchase decisions that are good for their profit line but in some ways bad for us. Compared to well-implemented public policy nudges, such private nudges are indeed less transparent and less democratically controlled. If public policy nudging is used to control or counteract many of those private influences, a nudge program could increase rather than diminish the democratic control we can exercise over our choice environments (Schmidt, 2017).

4.2 | Manipulation and dignity

Some critics worry that nudges are manipulative (Hausman & Welch, 2010; Mills, 2015; White, 2013; Wilkinson, 2013). Somewhat dramatically, Waldron (2014) thinks nudges are so manipulative they are an “an affront to human dignity.” Critics are particularly suspicious of paternalist nudges, as they seem to infantilize adults.

But what is manipulation and what makes it problematic? Consider two possible features of manipulation. Some instances of manipulation seem problematic as they influence people “behind their back” and “in the dark” instead of openly influencing them. But opacity is neither sufficient nor always necessary for manipulation (Sunstein, 2016, p. 104). We could imagine instances of manipulation that are done transparently or even announced, much like a hypnotist announcing their intentions at the beginning of their act. An oft-mentioned feature of manipulation is that it influences someone by bypassing their rational capacities or by exploiting their cognitive vulnerabilities (Blumenthal-Barby & Burroughs, 2012; Noggle, 2018). Now, if these tentative propositions about manipulation hold up, the manipulation objection seems to boil down to the objections discussed above, in particular those about transparency and rationality. Our contention thus is that to respond to the dignity and manipulation worry, one should best to attend to the earlier objections regarding autonomy, rationality, transparency, and dominating control.

4.3 | Illicit ends

A different worry is less about how nudging works but about the goals it promotes. The worry is that it might open up space for excessive paternalism and other illicit ends. A first concern is epistemological: How can policymakers know what is in the nudgee's best interest? A second concern is moral: The government should not impose their goals and values on citizens, particularly not in societies where individuals hold very different conceptions of the good (White, 2013).

In response, this objection hardly applies to nudging alone. Sure, there is the important question what aims are legitimate for public policy, but this holds for every policy tool. So, a defense of nudging, one might respond, does not require solving this problem. In addition, the epistemological concern applies to all attempts of (means) paternalism. What is more, one could argue that when people hold different ends and conceptions of the good, nudges have comparative advantages over other public policies like taxes, incentives, and prohibitions. As nudging preserves freedom of choice, we can resist or sidestep them and opt out.

But an objector might now reply that those seemingly desirable features also increase the risk of a slippery slope into too much paternalism and partisan political ends, for nudgers could always too easily turn around and argue that they keep freedom of choice intact (Rizzo & Whitman, 2009; Whitman, 2010).

The slippery slope argument is primarily empirical. From the armchair, it is unclear how likely the slope is to be slippery. Nonetheless, we should be mindful of the underlying worry. Are there potential mechanisms to filter out problematic goals and prevent a slippery slope? Here are some ideas.

First, in some cases, we can take the legitimacy of a policy goal for granted, such as achieving lower mortality rates in public health, higher tax compliance, or fewer traffic deaths. Here, the question is not whether to have public policy but what kind. Second, in other cases, proponents of nudging might want to defend the idea of *means paternalism* when people's ends are relatively clear, such as helping people lead longer lives and increase retirement savings (Sunstein, 2016, pp. 54–56). But the epistemological concern again may persist. While we can safely assume what people want in some cases (not die in traffic, avoid urine spillage, etc), other cases might raise stronger epistemological concerns. Finally, another institutional avenue to prevent a potential slippery slope is through transparency and democratic control. As mentioned above, like other policies, governmental nudges should be sufficiently transparent and subject to democratic accountability and control. Like with other policies, transparency and control should lower the risk of a problematic slippery slope by achieving a closer match between people's interests and the goals pursued through nudging (Schmidt, 2017, p. 413).

4.4 | Structural reform instead of nudging

A different worry focuses less on the techniques that nudges use but more on its use in real-world political and economic contexts. The worry is that nudges do little to address the underlying socioeconomic causes of many of the social ills we encounter. Obesity, for example, does not simply result from individuals making bad choices but has a complex biological and social etiology and comes with a social gradient. What is worse, some theorists worry that by shifting the focus towards individual choices and making individuals responsible for social ills—what Foucauldians call “responsibilization”—governments and society overall will become even less likely to address the real, underlying socioeconomic causes (Leggett, 2014, pp. 15–16; Peeters, 2017). Some have argued that the Conservative government in the UK, for example, have indeed sought to utilize nudging as part of a neoliberal agenda (Jones et al., 2011; Leggett, 2014, pp. 8–9).

Here are three responses.

The first, obvious response is that we do not have to choose between structural reform and nudging; we can do both. As a policy program, nudging is not meant to replace other policies and institutional reform programs. Moreover, proponents of nudging do not, or at least should not, advertise nudging as a panacea and we should not expect it to be the antidote to all structural and institutional problems.

Second, policymakers frequently find themselves in situations in which wider institutional reform is not feasible or would not get broad political support. In “nonideal” contexts, nudging then still has a role to play even for problems that would ideally be tackled through systematic and structural change. At the same time, the ideological critique does not apply to all nudges. Whatever one's political ideology, preventing traffic deaths or increasing tax compliance, for example, are desirable policy objectives.

Third, the nudge approach is not silent on at least some structural causes. As argued above, many of our choice environments have been intentionally shaped. For example, the nudge program implies that obesity is not merely a case of individuals making occasional bad decisions but that there often is a structural and systematic ill match. Nudging of course does not address all structural causes of our bad consumption habits. But public policy nudging moves our attention to one structural problem by foregrounding the structural mismatch between people's interests and the choice environments they encounter.

Finally, some proponents argue that nudging can even help address structural disadvantages. All humans operate with limited cognitive "bandwidth" and often in messy and complex environments. Yet people of lower socioeconomic status often have comparatively less cognitive bandwidth to spare and sometimes—out of necessity—rely on more short-term decision-making (Morton, 2017; Mullainathan & Shafir, 2013; Tully, 2019). This confluence of external disadvantage and decision-making can reinforce disadvantage. Nudge interventions designed to improve choice architecture for disadvantaged populations can thus contribute towards addressing—without of course solving—structural disadvantage (Tully, 2019).

5 | CONCLUSIONS AND FUTURE DIRECTIONS

In this article, we discussed some central arguments in the by now rich ethical debate around nudging. We first considered arguments that nudging promises cost-effective, evidence-driven, freedom-preserving, nonpartisan, and comparatively popular policies. When discussing arguments against nudging, we first grouped some of them along the different dimensions of autonomy.

First, some commentators worry that nudging undermines volitional autonomy and others that it undermines, or fails to respect, rational agency. In response, one might argue that the counterfactual to nudging often does not leave agents any more volitionally autonomous or rational. Moreover, some nudge policies might even be used to more closely align people's decisions to their ends and improve their decision-making, thereby strengthening volitional autonomy and rational agency.

Second, some worry that nudge programs in public policy might lead to dominating power relationships. In response, we highlighted that this risk will depend on *how* public policy nudges are pursued. If done democratically and transparently, nudging can meet the domination challenge.

Third, when discussing other worries—particularly manipulation and dignity—we argued that those are also best understood, and responded to, within discussions around personal autonomy.

Finally, we considered political worries that nudge programs might be used to pursue illicit ends and hinder wider structural reform. In response, we argued that with the right institutional framework, and in combination with other policies, nudge programs likely escape these worries.

Overall, our view is that no "knockdown" objection has appeared to date that should make us reject the nudge approach overall. At the same time, serious ethical concerns have emerged that should guide and inform discussions around whether and which particular nudge policies should be pursued, and, if so, how.

We now conclude with some general remarks on future directions in these debates.

First, while much of the discussion has rightly focused on tensions between the practice of nudging and important moral values, there is an increasing understanding that ethical concerns differ radically from practice to practice and from nudge to nudge. Instead of getting stuck in a stand-off between proponents and opponents making overgeneralizing claims ("nudging violates autonomy" versus "nudging promotes autonomy"), ethical analysis should proceed case by case (Engelen, 2019a; Sunstein, 2015a). The use of defaults in organ donation differs drastically from the design of urinals. Not only are the techniques employed and the heuristics triggered different, but more importantly, the goals and stakes diverge radically too—something any sensible ethical discussion must consider.

Moreover, while Thaler and Sunstein (2008, p. 252) presented nudging as "the real third way" that would break the deadlock between liberals and conservatives, we should recognize that nudging is not the panacea some hoped

it would be. Neither does it amount to a horror scenario of a world in which every single choice we make is being managed, as feared by (Waldron, 2014).

Notwithstanding a stronger focus on individual policies, the ethical debate so far will help us structure how we approach the salient moral considerations that bear on whether and how particular nudge policies should be designed and implemented. For example, discussions around transparency and democratic control can direct our attention to how nudge policies can be implemented and communicated. Discussions around rationality can urge upon us questions of what alternatives there could be or how nudging can be used to match up with individual decision-making capacities in ways that empower citizens as agents.

Second, at various stages, we touched upon the relationship between public policy nudging and private choice architecture, particularly done by private companies. Future research might foreground general ethical issues surrounding nudging across private and public spheres. But it might also zoom in on specific fields where choice architecture is being applied. In our view, a particularly fruitful and underresearched area is “digital choice architecture,” that is, the way data-driven nudges in smartphones, social networks, websites, and so on steer us in evermore precise ways (Alfano, Carter, & Cheong, 2018; Susser, Roessler, & Nissenbaum, 2019; Weinmann, Schneider, & Brocke, 2016; Yeung, 2017). Several features make this a particularly fertile ground for ethical inquiry. First, digital nudges are often not only inevitable but also effective, as companies can collect user data, perform A/B tests, use algorithms, and so on to shape user behavior. Second, smartphone apps and social media like Facebook and Twitter often systematically prey on our weaknesses for instant rewards and social comparison. Digital choice architecture often stands in the way of more fulfilling ways of directing our attention. Finally, the way such choice architecture works is far from transparent, as most technical details are not disclosed and—even when they are—often remain beyond the grasp of even the more tech-savvy consumers. Machine learning algorithms could even lead to “black boxes” in choice architecture where we do not even understand afterwards why an algorithm presented choices or content the way it did.

Clearly, digital environments provide fertile ground for effective and personalized choice architecture, without such choice architecture being transparent or democratically controlled to track our central interests. At the same time, our current approach to digital choice architecture arguably focuses too much on individuals while neglecting collective action. Individuals are typically expected themselves to bear the responsibility to deal with digital choice architecture and temptation. Many of the issues that emerged in ethical discussions around public policy nudging should thus provide an interesting entryway into the ethics of digital choice architecture and the measures most appropriate to improve it.

ORCID

Andreas T. Schmidt  <https://orcid.org/0000-0002-2323-2838>

ENDNOTE

¹ See Sunstein (2013) and Thaler and Sunstein (2008) for nudging in general, Lourenço, Ciriolo, Almeida, and Troussard (2016) for nudging in the EU, “Social and Behavioral Sciences Team; Annual Report” (2015) for nudging in the USA and Halpern (2015) in the UK, and “World Development Report (2015): Mind, Society, and Behavior” for behavioral policies in developing countries.

WORKS CITED

- Alfano, M., Carter, J. A., & Cheong, M. (2018). Technological seduction and self-radicalization. *Journal of the American Philosophical Association*, 4(3), 298–322. <https://doi.org/10.1017/apa.2018.27>
- Arno, A., & Thomas, S. (2016). The efficacy of nudge theory strategies in influencing adult dietary behaviour: A systematic review and meta-analysis. *BMC Public Health*, 16(1), 676. <https://doi.org/10.1186/s12889-016-3272-x>
- Barton, A., & Grüne-Yanoff, T. (2015). From libertarian paternalism to nudging—And beyond. *Review of Philosophy and Psychology*, 6(3), 341–359. <https://doi.org/10.1007/s13164-015-0268-x>

- Blumenthal-Barby, J. S., & Burroughs, H. (2012). Seeking better health care outcomes: The ethics of using the "nudge". *The American Journal of Bioethics*, 12(2), 1–10. <https://doi.org/10.1080/15265161.2011.634481>
- Bovens, L. (2009). The ethics of nudge. In T. Grüne-Yanoff & S. O. Hansson (Eds.), *Preference change: Approaches from philosophy, economics and psychology* (pp. 207–219). Berlin and New York: Springer Science & Business Media. https://doi.org/10.1007/978-90-481-2593-7_10
- Carlson, P. J., Park, E. S., & Andersen, C. K. (2009). Benefits of pavement markings: A renewed perspective based on recent and ongoing research. *Transportation Research Record*, 2107(1), 59–68. <https://doi.org/10.3141/2107-06>
- Christman, J. (1991). Autonomy and personal history. *Canadian Journal of Philosophy*, 21(1), 1–24. <https://doi.org/10.1080/00455091.1991.10717234>
- Conly, S. (2013). *Against autonomy: Justifying coercive paternalism*. Cambridge: Cambridge University Press.
- Engelen, B. (2019a). Ethical criteria for health-promoting nudges: A case-by-case analysis. *American Journal of Bioethics*, 19(5), 48–59. <https://doi.org/10.1080/15265161.2019.1588411>
- Engelen, B. (2019b). Nudging and rationality: What is there to worry? *Rationality and Society*, 31(2), 204–232. <https://doi.org/10.1177/1043463119846743>
- Engelen, B., Nys, T. Nudging and Autonomy: Analyzing and Alleviating the Worries. *Rev.Phil.Psych.* (2019). <https://doi.org/10.1007/s13164-019-00450-z>
- Fong, G. T., Hammond, D., & Hitchman, S. C. (2009). The impact of pictures on the effectiveness of tobacco warnings. *Bulletin of the World Health Organization*, 87(8), 640–643. <https://doi.org/10.2471/BLT.09.069575>
- Frankfurt, H. G. (1971). Freedom of the will and the concept of a person. *The Journal of Philosophy*, 68(1), 5–20. <https://doi.org/10.2307/2024717>
- Furedi, Frank. 2011. 'Defending Moral Autonomy against an Army of Nudgers'. *Spiked*. 20 January 2011. <http://www.spiked-online.com/newsite/article/self-help/10102>.
- Garnett, M. (2014). The autonomous life: A pure social view. *Australasian Journal of Philosophy*, 92(1), 143–158. <https://doi.org/10.1080/00048402.2013.765899>
- Gigerenzer, G., & Selten, R. (2002). *Bounded rationality: The adaptive toolbox*. Cambridge: MIT Press. <https://doi.org/10.7551/mitpress/1654.001.0001>
- Gigerenzer, Gerd, and Henry Brighton. 2009. 'Homo Heuristicus: Why Biased Minds Make Better Inferences'. *Topics in Cognitive Science* 1 (1): 107–43. <https://doi.org/10.1111/j.1756-8765.2008.01006.x>
- Grüne-Yanoff, T. (2012). Old wine in new casks: Libertarian paternalism still violates liberal principles. *Social Choice and Welfare*, 38(4), 635–645. <https://doi.org/10.1007/s00355-011-0636-0>
- Hagman, W., Andersson, D., Västfjäll, D., & Tinghög, G. (2015). Public views on policies involving nudges. *Review of Philosophy and Psychology*, 6(3), 439–453. <https://doi.org/10.1007/s13164-015-0263-2>
- Halpern, D. (2015). *Inside the nudge unit: How small changes can make a big difference*. London: Penguin.
- Hansen, P. G., & Jespersen, A. M. (2013). Nudge and the manipulation of choice: A framework for the responsible use of the nudge approach to behaviour change in public policy. *European Journal of Risk Regulation*, 4(3), 3–28.
- Hausman, D. M., & Welch, B. (2010). Debate: To nudge or not to nudge*. *Journal of Political Philosophy*, 18(1), 123–136. <https://doi.org/10.1111/j.1467-9760.2009.00351.x>
- Heilmann, C. (2014). Success conditions for nudges: A methodological critique of libertarian paternalism. *European Journal for Philosophy of Science*, 4(1), 75–94. <https://doi.org/10.1007/s13194-013-0076-z>
- Hertwig, R., & Grüne-Yanoff, T. (2017). Nudging and boosting: Steering or empowering good decisions. *Perspectives on Psychological Science*, 12(6), 973–986. <https://doi.org/10.1177/1745691617702496>
- Hummel, D., & Maedche, A. (2019). How effective is nudging? A quantitative review on the effect sizes and limits of empirical nudging studies. *Journal of Behavioral and Experimental Economics*, 80(June), 47–58. <https://doi.org/10.1016/j.socec.2019.03.005>
- Ivanković, V., & Engelen, B. (2019). Nudging, transparency, and watchfulness. *Social Theory and Practice*, 45(1), 43–74. <https://doi.org/10.5840/soctheorpract20191751>
- John, P. (2018). *How far to nudge? Assessing behavioural public policy*. Cheltenham: Edward Elgar Publishing.
- Jones, Rhys, Jessica Pykett, and Mark Whitehead. 2011. 'Governing temptation: Changing behaviour in an age of libertarian paternalism'. *Progress in Human Geography* 35 (4): 483–501. <http://dx.doi.org.ezproxy.princeton.edu/10.1177/0309132510385741>.
- Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.
- Leggett, W. (2014). The politics of behaviour change: Nudge, neoliberalism and the state. *Policy & Politics*, 42(1), 3–19. <https://doi.org/10.1332/030557312X655576>
- Lepenies, Robert, and Magdalena Małecka. 2019. 'The ethics of behavioural public policy'. In *The Routledge Handbook of Ethics and Public Policy*, edited by Annabelle Lever and Andrei Poama, 513–25. Abingdon: Routledge.
- Lourenço, J. S., Ciriolo, E., Almeida, S. R., & Troussard, X. (2016). Behavioural insights applied to policy: European report. In *EUR 27726 EN*. <https://doi.org/10.2760/903938>. Brussels, Belgium: European Commission Joint Research Centre.

- Lovett, F. (2010). *A general theory of domination and justice*. Oxford: Oxford University Press.
- MacKay, D., & Robinson, A. (2016). The ethics of organ donor registration policies: Nudges and respect for autonomy. *The American Journal of Bioethics*, 16(11), 3–12. <https://doi.org/10.1080/15265161.2016.1222007>
- Mills, C. (2015). The heteronomy of choice architecture. *Review of Philosophy and Psychology*, 6(3), 495–509. <https://doi.org/10.1007/s13164-015-0242-7>
- Morton, J. M. (2011). Toward an ecological theory of the norms of practical deliberation. *European Journal of Philosophy*, 19(4), 561–584. <https://doi.org/10.1111/j.1468-0378.2010.00400.x>
- Morton, J. M. (2017). Reasoning under scarcity. *Australasian Journal of Philosophy*, 95(3), 543–559.
- Mullainathan, Sendhil, and Eldar Shafir. 2013. *Scarcity: Why Having Too Little Means So Much*. New York: Henry Holt & Co.
- Niker, Fay. 2014. 'Nudge, Nudge? Privatizing Public Policy'. *Justice Everywhere* (blog). October 2, 2014. <http://justice-everywhere.org/?s=nudge>.
- Noggle, R. (2018). The ethics of manipulation. In E. N. Zalta (Ed.), Summer 2018 *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab: Stanford University. <https://plato.stanford.edu/archives/sum2018/entries/ethics-manipulation/>
- Nys, T. R. V., & Engelen, B. (2017). Judging nudging: Answering the manipulation objection. *Political Studies*, 65(1), 199–214. <https://doi.org/10.1177/0032321716629487>
- Oshana, M. (2006). *Personal autonomy in society*. Aldershot: Ashgate Publishing, Ltd.
- Peeters, R. (2017). *Manufacturing responsibility: The governmentality of behavioural power in social policies* (pp. 1–15). November: Social Policy and Society. <https://doi.org/10.1017/S147474641700046X>
- Petrescu, D. C., Hollands, G. J., Couturier, D.-L., Ng, Y.-L., & Marteau, T. M. (2016). Public acceptability in the UK and USA of nudging to reduce obesity: The example of reducing sugar-sweetened beverages consumption. *PLoS One*, 11(6), e0155995.
- Pettit, P. (1997). *Republicanism: A theory of freedom and government*. Oxford: Oxford University Press.
- Pettit, P. (2014). *Just freedom: A moral compass for a complex world*. New York: W. W. Norton & Company.
- Reisch, L. A., & Sunstein, C. R. (2016). Do Europeans like nudges? *Judgment and Decision Making*, 11(4), 310–325.
- Reisch, L. A., Sunstein, C. R., & Gwozdz, W. (2017). Beyond carrots and sticks: Europeans support health nudges. *Food Policy*, 69(May), 1–10. <https://doi.org/10.1016/j.foodpol.2017.01.007>
- Rithalia, A., McDaid, C., Suekarran, S., Myers, L., & Sowden, A. (2009). Impact of presumed consent for organ donation on donation rates: A systematic review. *BMJ*, 338(January), a3162. <https://doi.org/10.1136/bmj.a3162>
- Rizzo, M. J., & Whitman, D. G. (2009). Little brother is watching you: New paternalism on the slippery slopes. *Arizona Law Review*, 51, 685–739.
- Rebonato, Riccardo. 2014. 'A Critical Assessment of Libertarian Paternalism'. *Journal of Consumer Policy* 37 (3): 357–96. <https://doi.org/10.1007/s10603-014-9265-1>.
- Saghai, Y. (2013). Salvaging the concept of nudge. *Journal of Medical Ethics*, 39(8), 487–493. <https://doi.org/10.1136/medethics-2012-100727>
- Schmidt, A. T. (2017). The power to nudge. *American Political Science Review*, 111(2), 404–417. <https://doi.org/10.1017/S0003055417000028>
- Schmidt, A. T. (2018). Domination without inequality? Mutual domination, republicanism, and gun control. *Philosophy & Public Affairs*, 46(2), 175–206. <https://doi.org/10.1111/papa.12119>
- Schmidt, A. T. (2019). Getting real on rationality—Behavioral science, nudging, and public policy. *Ethics*, 129(4), 511–543. <https://doi.org/10.1086/702970>
- 'Social and Behavioral Sciences Team; Annual Report'. 2015. Executive Office of the President; National Science and Technology Council. https://www.whitehouse.gov/sites/default/files/microsites/ostp/sbst_2015_annual_report_final_9_14_15.pdf.
- Sunstein, C. R. (2013). *Simpler: The future of government*. New York: Simon and Schuster.
- Sunstein, C. R. (2015a). The ethics of nudging. *Yale Journal of Regulation*, 32, 413–450.
- Sunstein, C. R. (2015b). Nudges do not undermine human agency. *Journal of Consumer Policy*, 38(3), 207–210. <https://doi.org/10.1007/s10603-015-9289-1>
- Sunstein, C. R. (2016). *The ethics of influence: Government in the age of behavioral science*. New York: Cambridge University Press.
- Sunstein, C. R. (2019). *The cost-benefit revolution*. Cambridge, Mass: MIT Press.
- Susser, D., Roessler, B., & Nissenbaum, H. (2019). Technology, autonomy, and manipulation. *Internet Policy Review*, 8(2), 22.
- Taylor, C. (1979). What's wrong with negative liberty. In A. Ryan (Ed.), *The idea of freedom* (pp. 175–193). Oxford: Oxford University Press.
- Thaler, R. H., & Sunstein, C. (2008). *Nudge*. New Haven: Yale University Press.

- Todd, P. M., & Gigerenzer, G. (2012). Ecological rationality: The normative study of heuristics. In P. M. Todd, G. Gigerenzer, & ABC Research Group (Eds.), *Ecological rationality: Intelligence in the world* (pp. 487–497). New York: Oxford University Press.
- Todd, P. M., Gigerenzer, G., & ABC Research Group. (2012). *Ecological rationality: Intelligence in the world*. USA: Oxford University Press.
- Tully, Kendra. 2019. 'Odd bedfellows: How choice architecture can enhance autonomy and mitigate inequality'. *Behavioural Public Policy* online first (December): 1–19. <https://doi.org/10.1017/bpp.2019.36>.
- Vallier, K. (2016). On the inevitability of nudging. *Georgetown Journal of Law & Public Policy*, 14, 817–828.
- Waldron, J. (2014). It's all for your own good. *The New York Review of Books*, (October 9, 2014). <http://www.nybooks.com/articles/archives/2014/oct/09/cass-sunstein-its-all-your-own-good/>
- Weinmann, M., Schneider, C., & Brocke, J. v. (2016). Digital nudging. *Business & Information Systems Engineering*, 58(6), 433–436. <https://doi.org/10.1007/s12599-016-0453-1>
- White, M. D. (2013). *The manipulation of choice: Ethics and libertarian paternalism*. New York: Palgrave Macmillan.
- Whitman, D. G. (2010). The rise of the new paternalism. *Cato Unbound.*, (April 5, 2010). <http://www.cato-unbound.org/2010/04/05/glen-whitman/rise-new-paternalism>
- Wilkinson, T. M. (2013). Nudging and Manipulation. *Political Studies*, 61(2), 341–355. <https://doi.org/10.1111/j.1467-9248.2012.00974.x>
- Willis, L. E. (2013). When nudges fail: Slippery defaults. *University of Chicago law Review*, 80(3), 1155–1230.
- World Development Report. (2015). *Mind, society, and behavior*. 2015. Washington, DC (p. 2015). World Bank Group: World Bank. <https://openknowledge.worldbank.org/handle/10986/20597>
- Yeung, K. (2017). "Hypenudge": Big data as a mode of regulation by design. *Information, Communication & Society*, 20(1), 118–136. <https://doi.org/10.1080/1369118X.2016.1186713>

AUTHOR BIOGRAPHIES

Andreas T. Schmidt is an Assistant Professor of Political Philosophy at the Philosophy Faculty and the Centre for PPE in Groningen in the Netherlands. He holds a DPhil in Philosophy from the University of Oxford. His research is in political theory, ethics, and the philosophy of public policy, including topics such as freedom, equality, health policy, consequentialism, distributive justice/ethics, behavioral policies, longtermism, and animal ethics. See www.andreastschmidt.com for more.

Bart Engelen studied Philosophy at the KU Leuven, where he also did his PhD and worked as postdoctoral researcher. In 2014, he moved to Tilburg University (The Netherlands) and is now affiliated with the Tilburg Center for Moral Philosophy, Epistemology and Philosophy of Science (TiLPS). He teaches ethics, social, and political philosophy to students in philosophy, business economics, human resource management, and sociology. His research is situated on the borders between ethics, political philosophy (institutional design), and economics (rational choice theory). While he has published on conceptual and normative issues surrounding rationality, reciprocity, preferences, and (compulsory and open) voting, he has more recently focused on ethical and conceptual issues that arise with respect to nudging. Recent publications include "Ethical Criteria for Health-Promoting Nudges: A Case-by-Case Analysis" (*American Journal of Bioethics*, 2019), "Nudging and Autonomy: Analyzing and Alleviating the Worries" (*Review of Philosophy and Psychology*, 2019, together with Thomas Nys), and "Judging Nudging: Answering the Manipulation Objection" (*Political Studies*, 2017, together with Thomas Nys).

How to cite this article: Schmidt AT, Engelen B. The ethics of nudging: An overview. *Philosophy Compass*. 2020;15:e12658. <https://doi.org/10.1111/phc3.12658>