Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-025-04708-z OPEN

Individual autonomy and public deliberation in behavioral public policy

Nathanaël Colin-Jaeger[1] & Malte Dold[2]✉

The most successful concept in behavioral public policy (BPP) is nudging, which involves altering choice architecture to leverage people's biases and heuristics to promote welfare-improving behaviors. However, in recent years, nudging has faced criticism. This article addresses a specific critique: while nudges may enhance welfare, they often fail to promote autonomy. Several authors have raised this concern, yet there is no unified definition of autonomy in BPP. This article delves into the various meanings of autonomy in the BPP literature: freedom of choice, agency, and self-constitution. It focuses on autonomy as self-constitution, which acknowledges instrumental rationality but also considers substantive rationality, i.e., people's ability to reason about their goals, aspirations, and identities. The article explores epistemic, normative, and psychological challenges of autonomy as self-constitution and suggests that public deliberation in mini-publics could mitigate some of these challenges. Moreover, it emphasizes that an autonomy-centric BPP should shift its focus from reframing individual choice situations (i-frame interventions) to enabling public deliberation about institutional choices (s-frame interventions).

[1] American University of Paris, Paris, France. [2] Pomona College, Claremont, USA. ✉email: malte.dold@pomona.edu

Introduction

A core contributor to Behavioral Public Policy's (BPP) ongoing success is the idea of *nudging* which is explicitly built on behavioral insights to help people overcome problems of bounded rationality or willpower (Thaler and Sunstein, 2008). Its proponents describe a *nudge* as "any aspect of the choice architecture that alters people's behavior predictably without forbidding any options or significantly changing their economic incentives." (Thaler and Sunstein, 2008, p. 6). Nudges straddle the line between libertarianism and paternalism. They aim to uphold individual autonomy by ensuring that individuals can easily opt for the non-nudged option if they so wish. Simultaneously, they adopt a paternalistic stance, as they involve choice architects guiding individuals towards welfare-improving outcomes by reshaping the decision-making environment. Thaler and Sunstein provide a limited discussion of what constitutes individual welfare. Yet, they emphasize the importance of respecting the subjective goals of individuals when they say that nudges aim to "make choosers better off, *as judged by themselves*" (Thaler and Sunstein, 2008, p. 5).

As is the case with other successful policy ideas, the idea of nudging has been defended and criticized.[1] This article does not rehash this debate but focuses on a strand in the BPP literature that has advocated for a shift away from the narrow welfarist focus of nudging toward a more autonomy-oriented interpretation of BPP (Banerjee et al., 2024; Hargreaves Heap, 2023; Dold and Lewis, 2023). This shift has been motivated by the observation that nudges often leverage people's cognitive biases to prompt behavioral shifts. Instead of providing reasons for choosing the nudged option, nudges exploit the same decision-making flaws that led to the initial biased choice. Altering the default in retirement savings plans is a prime example (Beshears et al., 2009). Here, the nudge doesn't make people more competent in financial matters or empower them to reflect on their preferences; rather, the nudge exploits people's tendency to go with the status quo, thereby steering them toward higher savings rates. The individual being nudged often remains unaware of the intervention's purpose. Choice architects reshape the decision environment from the top down, aiming to steer individuals towards options deemed best by the architect. Consequently, nudges often "leave citizens 'in the dark', incapable of internalizing and owning the process of behavior change" (Banerjee et al., 2024, p. 1).

While the autonomy-oriented strand in BPP is united in this critique of the nudging approach, it has not come up with a unified understanding of what constitutes autonomy (Vugts et al., 2020). In Section Autonomy in BPP: Freedom of choice, agency, and self-constitution, we discuss the various understandings of autonomy proposed in the literature. We focus on a more expansive understanding of *autonomy as self-constitution* and the idea that behavioral insights can contribute to a discussion of interventions that enhance people's capability to reflect on their identity-shaping choices. A core goal of BPP that focuses on autonomy as self-constitution is to help people become aware of situational and social factors shaping their belief and preference formation processes and, in doing so, increase their capacity to form intentions self-reflectively and act freely on them. We admit in Section The threefold challenge of autonomy as self-constitution that this proposal for a 'thicker' notion of autonomy in BPP is not without its problems. In particular, it faces *epistemic, normative, and psychological challenges*. *Epistemically*, a concern is the question of whether policy designers can gain knowledge of the kind of identity formation processes individuals favor. *Normatively*, a concern is whether an autonomy-oriented BPP can avoid falling back onto a form of autonomy paternalism that might reflect the goals of academically trained policy experts but miss people's perspectives. *Psychologically*, a concern is whether individuals have the cognitive means and motivation to engage in transpositional exercises to form their identities in a self-reflective manner.

In Section Public Deliberation: Autonomy Through Self-Government, we discuss ways in which public deliberation can help mitigate these challenges. This builds on the pioneering work of John et al. (2011), who were among the first to integrate and evaluate various deliberative mechanisms as tools for behavioral public policy, based on what they call a 'think' strategy. Their central idea is that citizens should engage in deliberation and set their own priorities, fostering a process of civic and democratic renewal. Public deliberation builds on the concept of "talking as a decision procedure" (Schauer, 1999). While the exchange and pooling of information are essential to the deliberative process, its core lies in fostering discussion, disputation, and the transformation of participants' views (Elster, 1986/2010). If properly designed, public deliberation can enhance participants' reasoning abilities by exposing them to a diverse set of beliefs and evaluative standards, thereby making them aware of their own positionality (Sen, 2009; Mercier and Sperber, 2011; Dold, 2024). In essence, public deliberation can enhance individual autonomy by fostering trans-positional reasoning, thereby supporting individuals in their processes of self-constitution. This helps tackle the *psychological challenge*. Public deliberation can take various forms, such as mini-publics, citizen juries, deliberative polling, or consensus conferences. In this article, we do not advocate for a specific institutional form of public deliberation but rather underscore its constructive aspects to also tackle the *epistemic* and *normative challenges* mentioned before. By involving citizens in the policy design process, public deliberation alleviates the epistemic burden on experts in the context of BPP. Furthermore, by relying on the buy-in of affected citizens, it addresses the normative concern of paternalism. We acknowledge that public deliberation in the form of citizen participation in public policy processes encounters counterarguments. We draw upon the empirical literature on public deliberation to address some of them and advocate for a well-designed deliberative process.

In a nutshell, the argument of the paper is as follows: We argue that self-constitution is a valuable aspect of what it means to exercise one's autonomy, as it empowers individuals to choose, through a variety of means, which individual they want to become. If fostering individual self-constitution is a legitimate political goal, public deliberation is an institutional response that helps mitigate the challenges that this conception of autonomy introduces to the policy-making process. Overall, our article invites scholars in BPP to move beyond a narrow focus on nudge or boost interventions and recognize that addressing autonomy requires deeper engagement with the broader socio-political institutions that shape individuals' capacity for self-constitution.

Autonomy in BPP: Freedom of choice, agency, and self-constitution

While policy proposals in BPP typically aim at increasing people's welfare—e.g., by nudging them to save more or eat fewer fatty foods—they typically also rely to a large degree on the idea of autonomy: the consumer is supposed to make the ultimate decision and is free to ignore the nudge.[2] Both the proponents of nudging, such as Thaler and Sunstein (2008), and contemporary critics who favor different behaviorally-informed interventions, such as Banerjee et al. (2024), are advocating for policies that preserve or promote personal autonomy. However, autonomy is a broad church. Following Vugts et al. (2020), one can distinguish at least three conceptions of autonomy in BPP: autonomy as

freedom of choice, autonomy as *agency*, and autonomy as *self-constitution*. Most debates in BPP revolve around the first and second conception, while we argue in this article that the third conception ought to be properly taken into account. We argue that understanding autonomy as self-constitution requires BPP scholars to focus on the institutional and political conditions for autonomy (s-frame), and not only on the individual conditions for welfare-improving choices (i-frame).[3] If autonomy as self-constitution is valuable, an institutional turn in BPP is warranted. But before discussing institutional implications, let us first depict the three conceptions of autonomy.

The first conception of autonomy is *freedom of choice*. This is the main conception of autonomy that we find in BPP, in particular in the early discussions of nudging (Thaler and Sunstein, 2008) and its subsequent refinements in response to critics (e.g., Sunstein, 2015, 2016, 2020). The main idea is that behavioral policies should respect individuals' freedom to choose among different options, such that even when people are nudged, they can switch at no or a low cost to another option. In other words, nudges do not take options off the table or significantly alter their relative prices. Nudges and other interventions are seen as legitimate if they are not coercive and people are left free to go their own way.

The typical cases where nudges are invoked are cases of *weakness of the will*—when individuals wish to lose weight but succumb to their urgent needs when they see chocolate at the cafeteria cashier—or *lack of information*—when they have to choose a retirement plan but get lost in the complexity of the choice menu. In these cases, nudges can help individuals to act according to their goals (in this case, losing weight or saving effectively for retirement). Nudges include opting people automatically into employer-matched retirement savings plans or rearranging items in the cafeteria to make them less visible, thus helping people to choose healthy options. The hope is in both cases that individuals would accept the nudge since it steers them toward their 'true' preferences (losing weight, having a good retirement plan) against their seemingly biased preferences (succumbing to temptation, saving less).[4]

Kuyer and Gordijn (2023) argue that most discussions on the ethics of nudging (and BPP more broadly) occur within this conception of autonomy as freedom of choice. Interestingly, it is not just the proponents of nudging who subscribe to this understanding of autonomy but also some of its critics. For example, Sugden (2018) criticizes Thaler and Sunstein's approach based on the argument that they do not provide a realistic psychological model of what constitutes 'true' preferences and, in doing so, do not take the inherent context-dependent nature of preferences seriously. Nevertheless, Sugden argues that autonomy should be understood as the freedom to choose among options, such that autonomy is increased when the opportunity set is enlarged.

The second conception of autonomy in BPP acknowledges that freedom of choice is undoubtedly valuable (in the sense that people care about it), but not self-standing. It should be complemented with a focus on agency understood as decision competence (Dold, 2023). This view attracted considerable attention over the past years, particularly by proponents of the boosting approach (Grüne-Yanoff and Hertwig, 2016; Hertwig and Grüne-Yanoff, 2017; Banerjee et al., 2024). One idea underlying boosting is that BPP should not just be concerned with the number of opportunities people are confronted with, but also their capacity to choose competently among them (Dold and Rizzo, 2021). In this second conception of *autonomy as agency*, autonomy requires not only having options and being free from coercion but also having the internal capacity to reflect on what options are best to achieve a given personal goal (Vugts et al., 2020). A further difference with the first conception of autonomy is that it concerns agents' internal cognitive processes, as it requires BPP to increase people's mental capacities to reason instrumentally and choose competently (McKenzie et al., 2018; Dold and Lewis, 2023).

Proponents of this notion of autonomy criticize nudges for violating autonomy as agency because they (largely) target or exploit shallow cognitive processes – the fast, automatic, and intuitive type 1 processes (Kahneman, 2011) – to alter people's behavior but do not try to improve their cognitive processes by appealing to rational and effortful type 2 thinking (Saghai, 2013). If this is true, then nudges largely fail to treat people as reasoning subjects, diminish the control over their choice process, and ultimately undermine their agency (Hausman and Welch, 2010; Lanzing, 2019; Noggle, 2018). According to proponents of the second notion of autonomy, specific BPPs, such as *boosts* or *debiasing*, should help people choose the 'right' means for a given end by helping them "build cognitive capability and motivation through the regular activation of reflective processes" (Banerjee et al., 2024).

Boosts modify people's behavior by making them more competent regarding a specific problem, typically by teaching them the application of simple decision rules (Grüne-Yanoff and Hertwig, 2016; Grüne-Yanoff, 2021). They consist, for instance, of special training regarding issues individuals face: self-control training to avoid procrastination or training in fast-and-frugal decision trees to break down complex decisions, e.g., in the medical context (Gigerenzer, 2024). Boosts encompass domain-specific training (e.g., how to exercise and remain motivated) or general skills (e.g., statistical literacy for various decisions). Boosts do not bypass people's system 2 thinking because training ought to be accepted and actively chosen by individuals. Boosted individuals select themselves into training activities that enable them to achieve their goals (e.g., becoming more reliable, more statistically trained, etc.).

Debiasing is a second possibility to increase autonomy as agency, consisting of procedures not to make biases disappear – which is probably impossible – but to make them more explicit (Banerjee et al. 2024). A typical example is food labeling and mixed framing. A brand could sell their food with a 75% fat-free label, which could be accompanied by a 25% fat label, thereby helping people to make better decisions. More expansive debiasing techniques work in institutional or organizational settings where time-costly interventions, such as role-play, are available. In a role-play scenario, different positions on a topic can be predetermined to avoid false unanimity or fallacies. If some individuals are required to play a pessimistic role on a project, this can help counterbalance optimism bias. These techniques can also be attainable for individuals. For example, when confronted with an important career choice, people can write a letter to their future selves (Chishima and Wilson, 2021). In these situations, going beyond their current self and taking the time to make reasons explicit allows people to realize potential biases in their intuitive approach to a decision problem. Enhancing autonomy via boosts and debiasing techniques goes beyond the aforementioned freedom of choice view of autonomy. However, it still regards the primary objective of BPP as bolstering individuals' *instrumental reasoning* – namely, selecting the most appropriate means to achieve their given ends. Within this context, the biggest threat to instrumental reasoning is subtle manipulation (e.g., in the form of sublime advertisements) that leads people to choose suboptimal means for their given ends (Vugts et al., 2020).

We propose to go a step further defending a third account of autonomy as self-constitution. According to the philosopher Christine Korsgaard (2009), self-constitution requires conceiving yourself as the cause of your actions. Korsgaard (2009, p. xi) claims: "For an action is a movement attributable to an agent as

its author, and that means that whenever you choose an action—whenever you take control of your own movements—you are constituting yourself as the author of that action, and so you are deciding who to be." Seeing your action as emanating from yourself means recognizing yourself as a cause rather than the action as the production of something *supposedly within me* (such as a counterfactual 'true' preference) or *in work through me* (such as an external, manipulative influence). BPP interventions like nudges typically target some given desirable goals but do not necessarily respect agency as self-constitution; they exploit my biases and in this sense *work through me*. Moreover, assuming hypothetical, counterfactual preferences I might not have right now but will be nudged towards, nudges target something *supposedly within me* without necessarily recognizing myself as a whole (Korsgaard, 2009, 18).

In contrast, *autonomy as self-constitution* entails the production of oneself through decisions where the chooser themselves have a sense that they themselves are ultimately the cause of their own actions. We shape our identity creatively through *active* choices. In this sense, autonomy does not only entail the capacity to choose what means we want to pursue for given ends (*instrumental reasoning*), but also the capacity to decide which goals we want to pursue in the first place and which person we want to become over time (*substantive reasoning*). This implies that individuals have the capacity to reflect on their goals and evaluate their desirability. Self-constitution is achieved when individuals can recognize themselves as the authors of their own actions. This aspect is captured by Korsgaard's constitutional model: "What makes an action *mine*, in the special way that an action is *mine*, rather than something that just *happens* in me? That it issues from my constitution, rather than from some force at work within me; that it is expressive of a law I give to myself, rather than a law imposed upon me from without." (Korsgaard, 2009, 160). Self-constitution is obtained when such recognition is possible, even when the action is retrospectively seen as erroneous ("In the end, it was not the right thing to do, but it was my choice"), such that individuals may recognize themselves in the process of choosing and acting.[5]

This notion of autonomy does not imply that individuals need to be free from all external influences or to choose every action consciously to be autonomous. Agents can recruit routines through conscious intentions, leading to non-conscious behavior. Consequently, self-constitution can be compatible boosts, provided simple heuristics are chosen by individuals as tools to free up cognitive resources or support the achievement of reasoned goals, i.e., when they result from a prior intentional and deliberate decision. Sociology and social psychology have taught us that individuals are not free from external influences when making choices. Yet, self-constitution means that people can reflect on processes of preference and belief formation and the ways they are affected by social and psychological influences. Self-constitution goes beyond merely acknowledging a higher-order preference (e.g., "I want to lose weight but lack motivation"). It encompasses the ability to critically examine and revise these higher-order preferences through reasoning. In this regard, the ideas of indoctrination or mental conditioning are seen as threats to autonomy: they undermine people's capacity to be the authors of their own lives since they make it difficult for them to be aware of where their preferences and beliefs come from (Vugts et al., 2020; Dold, 2024).

Ultimately, the citizen should be the final arbiter of the merits of autonomy as self-constitution as a goal in BPP. Nevertheless, at least three key reasons speak for its significance in policy discourse. First, experimental evidence suggests that individuals put an intrinsic value on personal control and active choices, above and beyond welfare considerations (Bartling et al., 2014; Bobadilla-Suarez et al., 2017). Second, psychological research demonstrates that a sense of self-determination is instrumentally valuable, leading to positive outcomes such as creativity, cognitive flexibility, sustained behavior change, and increased subjective well-being (Deci & Ryan, 1987). Third, from a normative perspective, focusing on individual autonomy as self-constitution in public policy aligns with value pluralism. It respects the diversity of individual values and preferences without dictating specific choices. Instead, it emphasizes the importance of developing the capabilities necessary to become aware of contextual influences and make self-determined decisions (Dold and Lewis, 2023).

In the existing BPP toolkit, autonomy as self-constitution may be addressed best with the idea of nudge+ (Banerjee and John, 2021). Nudge+ mixes the idea of traditional Thaler-Sunstein nudges with deliberation about their use and the goals of citizens. Conscious reflection is used to overcome the potentially harmful effects of nudges (e.g., reactance or subtle manipulation). For Banerjee and John, nudge+ can take many forms depending on the temporal coupling of nudges with reflection (simultaneous or sequential). Deliberation plays an essential role for nudge +: it allows people to reflect not only on an institutional intervention but also on their preferences and what nudging means for their identity formation processes. Other BPP interventions that go beyond the idea of nudge+ would be s-frame institutional reforms that expose people to multiple viewpoints and challenge their preconceived notions on certain topics through *education*, *experiments in living*, and various channels of *public deliberation* (Dold, 2024). *Education* plays a crucial role in contributing to discussions about the core elements of our identity, but also in understanding that "diversities can take many distinct forms and that we have to use our reasoning to decide how to see ourselves." (Sen, 2003). The concept of *experiments in living* (Mill, 1859/2003) highlights the significance of encountering diverse life experiences. This can be achieved either through unintentional exposure to a variety of lifestyles or by deliberately immersing oneself in different cultures, practices, and political systems. Due to space constraints, this article does not delve deeper into education and experiments in living as important means to foster autonomy as self-constitution but will instead focus on the central role of public deliberation in the next section.[6] Table 1 gives an

Table 1 Conceptions of autonomy in BPP.

	Autonomy I ***Freedom of Choice***	**Autonomy II** ***Agency***	**Autonomy III** ***Self-Constitution***
Main Target	Behavioral Outcomes	Instrumental Reasoning (IR)	Substantive Reasoning (SR)
Task of BPP	Steer people's behavior in a certain predefined welfarist direction	Increase competence, i.e., capacity to choose means m_i from $\{m_1 \ldots m_n\}$ for given goal x	Increase agentic capability, i.e., ability to choose goal x_i from $\{x_1 \ldots x_n\}$ that is self-endorsed and authentic
Threats	Coercion	Subtle manipulation	Indoctrination and mental conditioning
Policy	Nudges	Boosts, debiasing	Nudge +, public deliberation, experiments in living
Level	i-frame intervention	i-frame intervention	i- and s-frame interventions

overview of the dimensions of the three conceptions of autonomy discussed so far.

We want to emphasize that interventions fostering Autonomy II, such as boosts, can contribute to the realization of Autonomy III. Boosts can free up cognitive resources and teach competencies that enable individuals to address questions of self-constitution more effectively. In doing so, boosts can be seen as part of a policy mix in a responsive state, one that is supported by self-constituting individuals. In other words, boosts can be endorsed by self-constitutive individuals as ways to realize reasoned ends.

The threefold challenge of autonomy as self-constitution

Autonomy as self-constitution means that individuals shape themselves through choices rather than merely satisfying some given goals. The idea of self-constitution is appealing to those who contend that emancipation and experimentation are valuable components of a fulfilled life (Dold and Lewis, 2023). However, in this conception of autonomy it might not be straightforward for external observers to gain knowledge of the type of social environments (e.g., pluralistic-open vs. homogenous-closed) individuals favor for their identity formation processes. Moreover, this conception encounters the problem of justifying what kinds of identity formation processes should be supported and for whom; autonomy as self-constitution risks reintroducing social planner-like conceptions of the good life that might be normatively problematic. Finally, even if these two challenges can be resolved, individuals might be overwhelmed by the cognitive demands of substantive reasoning and the idea of exposing their identity and goals to critical scrutiny. Consequently, there is at least a threefold challenge of using autonomy as self-constitution in BPP:

(i) *Epistemic challenge*: How can policymakers gain knowledge of the social environments conducive to the identities individuals aspire to hold?
(ii) *Normative challenge*: How can policymakers avoid reverting to paternalism when fostering certain processes of identity formation over others?
(iii) *Psychological challenge:* Do individuals have the cognitive means to engage in transpositional exercises to form their identities in a self-reflective manner?

Considering these challenges, it appears that an autonomy-oriented BPP encounters similar issues as the standard welfarist justification for BPP. In the latter case, it is problematic if people are nudged into behavioral outcomes that do not respect individuals' actual choices. In the former case, guiding the process of self-constitution might be equally problematic for the same reason: direct interventions might be illegitimate, and meta-interventions on the social environments that provide the background against which individuals form their identities can be illegitimate because they unduly steer or restrict the range of possible identities an agent may have access to. The process view of the self (which is inherent to autonomy as self-constitution) presents a specific challenge: if neither the policymaker's standpoint nor the individual standpoint at particular moments in time seems relevant, from which standpoint may we overcome this problem? Neither the "I" perspective of the chooser (due to context-dependence and cognitive limitations) nor the "She" perspective of the policymaker (due to the epistemic and normative problems stated above) offers a satisfying solution.

A promising attempt to tackle the epistemic and the normative challenges has been made by Lecouteux and Mitroutchev (2023) in an article defending the "view from manywhere", which is deemed normatively superior to the "view from nowhere" (the "She" perspective of the planner) and the "view from somewhere" (the "I" perspective of the choosing individuals). The view from manywhere is based on an intra-subjective dialogical account of the different identities of a given individual. Following Lecouteux and Mitroutchev (2023), an individual may generate a set of context-dependent identities with possibly different and competing interests (think of a person's identity *A'* in the context when she is drinking during the evening and her identity *A"* in the context of a sober but slightly hung-over context the next morning). Lecouteux and Mitrouchev argue that what should have normative status is not the choice of one identity over the others (e.g., *A"* over *A'*), but what we may call the deliberative process that generates the underlying preferences through an open-ended confrontation of interests and perspectives.

We sympathize with this approach, as the idea of achieving a view from manywhere has a long tradition in liberal thought from Adam Smith's idea of the *impartial spectator* (Smith, 1859) to Amartya Sen's account of *transpositional reasoning* (Sen, 2009). Nevertheless, we have two fundamental concerns with the approach – at least when it is understood narrowly. First, it fails to address the aforementioned *psychological challenge* because it imposes a heavy cognitive burden on individuals, requiring them to undertake the complex deliberative process of reconciling different perspectives within their multiple identities on their own. Second, even if individuals are psychologically capable of generating a view from manywhere, it can still be significantly constrained by imagination, experience, social networks, and other factors that may hinder individuals from achieving the normative goal of comprehending a broader range of influences on their behavior and preference formation. The Lecouteux-Mitrouchev approach remains mostly at the individual level as an inner deliberation, while self-constitution does not limit itself to strictly individual deliberation; in fact, self-constitution occurs reflexively within social and institutional contexts. In the next section, we argue that public deliberation helps alleviate these concerns.

Public Deliberation: Autonomy Through Self-Government

By public deliberation we refer to the idea that "people come together, based on equal status and mutual respect, to discuss political issues they face and, based on those discussions, decide on the policies that will affect their lives" (Bächtiger et al., 2018, p. 2). Public deliberation theorists emphasize that deliberation is above all a mode of *self-governance*; they consider it the privileged mode of public decision-making for members of a society who value equality and justifiability.[7] Traditionally, public deliberation in a representative democracy manifests itself through discussions held in assemblies or within judicial proceedings. Yet, recent literature on deliberation emphasizes alternative, more experimental channels of deliberation, such as *mini-publics* (Niemeyer, 2011; Fishkin et al., 2024; Landemore, 2020). Mini-publics can take different forms such as *citizen juries* (randomly selected citizens who deliberate on a specific issue, e.g., how to make the city more inclusive), *deliberative polling* (a statistically representative group of citizens randomly selected to participate in a moderated discussion on particular issues and propose recommendations, e.g., citizens conventions like the French Citizen Convention for Climate), or *consensus convention* (when a group of citizens and experts meet to discuss a particular issue to reach a consensus). In this article, we do not advocate for a specific institutional form of mini-publics. However, we want to note that public deliberation can and should happen in diverse institutional forms and at various levels (federal, state, and local). Moreover, we emphasize the use of random, statistically representative sampling to prevent public deliberation from being dominated by more educated individuals or those with more free time, who might otherwise self-select into public forums. Representative

sampling can be enhanced by offering monetary compensation to participants, which helps encourage the participation of socio-economically marginalized groups that are typically difficult to engage.

A significant challenge of current BPPs, such as nudges and boosts, is that they are designed by experts who are supposed to nudge or boost people in specific directions (based on what experts imagine to be in the people's interest – be it welfare or autonomy). This implies that experts need to know a lot about people's preferences, including possible trade-offs, such as between people's welfare and autonomy concerns. Nudges might increase welfare through automatization and unconscious procedures and boost interventions might enhance people's autonomy by teaching them simple heuristics that improve the quality of their instrumental reasoning. But how can BPP experts gain knowledge about such trade-offs and avoid falling back onto a form of welfare or autonomy paternalism? We argue that deliberation and the inclusion of citizens in designing the policies can help meet these concerns.

There are three core functions of deliberative mechanisms displayed in the empirical literature (Niemeyer et al., 2023): the transparency and accountability function, the reduction of cognitive burden function, and the enhancement of cognitive abilities function. The first function of deliberation is to increase transparency and accountability which allows us to take up the *normative challenge* raised above. The reason-giving requirement of deliberation is what allows communication between different individuals with different perspectives and interests (Habermas, 1996; Gutmann, Thompson (1997); Bohman, 2004; Landemore, 2020). This ties into the second function which is the *reduction of the epistemic burden* on behalf of the policymakers because it allows individuals to participate in the policy-making process and signal the trade-offs they want to accomplish. When deliberative forums produce a set of concrete and actionable policies, the epistemic burden for policymakers is decreased: they can look at the outcomes of deliberative processes and do not need to infer values or aspirations from observed market behavior or noisy surveys. This allows us to take up the *epistemic challenge* raised above.[8] The third crucial function is that participation in deliberative forums *enhances people's reasoning capacities*. Deliberation often has a transformative effect on people's beliefs due to the discovery of information as well as on their evaluative standards due to people's exposure to other people's views and values (Knight and Johnson, 2014); by recognizing their own individual positionality people broaden their views and become aware of factors that influence their substantive reasoning (Habermas, 1996). In this sense, people do not only learn new information based on the diversity of public beliefs but are led to develop new preferences and values based on the diversity of evaluative standards.[9] Argumentation makes us aware of hidden assumptions and influences and can thus modify our views. If we take these insights seriously, we see how public deliberation answers better than individual deliberation the *psychological challenge* mentioned in the previous section. The empirical literature on public deliberation suggests indeed that participation in public forums such as mini-publics has a transformative effect that facilitates the *view from manywhere*, enabling a broader transpositional view on matters that can contribute to more informed ways of self-constitution and individual identity formation (Niemeyer, 2011; Niemeyer et al. 2023). Yet, deliberative arenas pose new problems for BPP. They may shift the locus of BPP from individual reasoning to collective discussion (*Counterargument I*). Moreover, collective deliberation may not necessarily help individuals improve their cognitive abilities and make better decisions for themselves (*Counterargument II*).

Counterargument I. Why should BPP instruments transition from individual reasoning to public deliberation? Isn't BPP primarily concerned with *individual* autonomy and welfare? To counter these points, we wish to underscore that BPP does not shape individuals' choice environments in isolation but has also consequences for the structure of the social environment within which people interact and choices take place. Moreover, BPP is public and concerns individuals as citizens as much as private individuals. Inner deliberation, highlighted by Lecouteux and Mitroutchev (2023), may be insufficient to assess BPP's legitimacy and acceptability. This does not imply that BPP should be 'holistic', i.e., solely concerned with the welfare of the society as a whole. It should instead acknowledge that it impacts the relations between individuals and their interdependent identity formation processes as much as their ability to reason about ends or achieve a given end competently. Thus, we contend that public deliberation is a crucial component of a conception of autonomy as self-constitution in BPP. The process view of autonomy as self-constitution relies on the notion that for an intervention to be legitimate, it should align with individuals' values or aspirations. To achieve autonomy as self-constitution, BPP can help people navigate the process of identity formation in a transpositional manner, which cannot be dictated from the outside by a social planner or determined strictly individually. Instead, it should, at least in some cases, involve a social process of identity formation, i.e., public deliberation.

Public deliberation possesses the same advantages as the multiple identities model of Lecouteux and Mitroutchev (2023) but with real individuals representing different framings of a problem at the same time. Instead of individual deliberation, where one individual weighs various reasons by imagining different current and future versions of herself, it is collective deliberation among many individuals with different perspectives that fosters multiple-frame reasoning (Bermudez, 2020). The relation between individual deliberation and collective deliberation is not disjunctive. Individual reasoning benefits from the confrontation of views in public deliberation (Colin-Jaeger et al., 2022). It also helps answer the normative question: how can we identify a locus of legitimacy if neither the "I" perspective nor the "She" perspective is satisfying? Thus, the appeal of public deliberation lies in its potential to replace an expert-driven agenda and democratize BPP.

Counterargument II. How can one make sure that deliberation transforms individual views in the "right" direction, i.e., that it helps them transcend their positionality and avoid producing entrenched or parochial preferences and beliefs? Rather than enhancing autonomy by making choices more deliberate, it can create new pervasive influences and forms of social control. Many critics argue that individuals are poor reasoners, and deliberation may produce adverse effects. The criticisms are based on the assertion that individuals often possess low levels of information (Lippmann, 1925; Brennan, 2016), the observation of adverse effects of collective deliberation on human intelligence (Sunstein, 2002), or the incapacity of collective deliberation to improve individual decision-making (Müller, 2019). These results are not surprising given the number of cognitive biases and reasoning errors discussed in psychology and behavioral economics (Kahneman, 2011). In addition, Banerjee et al. (2024) recognize the possibility of intentional manipulation of deliberative assemblies by third parties. Considering these concerns, it seems unclear at best whether deliberation is the appropriate route to enhancing individual reasoning. Nevertheless, recent empirical studies on mini-publics offer good arguments to counter the claim that deliberation cannot, in principle, justify the positive

transformative claim made above.[10] What the critics show us is that attention should be paid to the *context* and *design* of deliberation. Two primary arguments can be formulated in response to the critics of deliberation.

First, most empirical works casting doubts on *public* deliberation studies *individual* reasoning: "Much political science survey research demonstrates citizen incompetence in terms of solitary reasoning. Many pertinent psychological experiments involve decontextualized tasks with no interaction and no supportive environment providing participants with adequate information. But deliberation involves reasoning together, not individually." (Niemeyer et al., 2023, p. 2). As Mercier and Sperber (2011, 2017) argue, human reasoning – understood as a deliberate process, where one or more premises lead to a conclusion – is best when exercised in small teams or groups with flat hierarchies. One possible explanation developed by Mercier and Sperber (2017) is that human reasoning evolved in groups and was selected as a communication tool rather than an individual task. Group reasoning displays better results than individual reasoning on most issues, including bias detection, information search, and even depolarization when people are facing a non-abstract situation of problem-solving. Group reasoning can sometimes even separate the production of arguments from their evaluation, which leads to better handling of confirmation biases (Mercier and Sperber, 2011, p. 63).[11]

Second, deliberation not only performs better in addressing individual biases, but it can also be enhanced through proper design. Niemeyer et al. (2023) emphasize the significance of *group building* and its impact on the quality and effectiveness of deliberation, extending beyond the conventional design elements of deliberative forums (such as diverse expertise, participant representativeness, process facilitation, etc.). Group building encompasses the idea that members of a group debate about the rules and roles within the group before the actual debate about substantive issues starts. Groups that display high levels of group building, particularly when they can develop their norms and rules of decision-making beforehand, share more information, are less polarized, and face complex issues more efficiently.

For an example where public deliberation could help increase individual autonomy, consider potential BPPs regarding the consumption of addictive products, such as tobacco, alcohol, or fatty foods. The BPP literature has suggested many ingenious ways to counter the self-control issues involved – from nudges (Thaler and Sunstein, 2008) to sin taxes (O'Donoghue and Rabin, 2006) to motivational boosts (Hertwig and Grüne-Yanoff, 2017). These suggestions have in common that they involve an expert who designs the nudge, calibrates the sin tax, and teaches the boost from a "She" perspective. Moreover, in the nudge and sin tax case, the affected individuals are not forming deliberate intentions themselves regarding how much tobacco, alcohol, or fatty foods they want to consume (in other words, their autonomy as self-constitution is sidestepped). In contrast, if the issue of consumption of addictive products is opened to public deliberative discourse, we allow the individual to be exposed to multiple positions on the issue. Relevant considerations to inform the discussion might include scientific findings on the addictive properties and physical consequences of specific goods and substances, as well as research on the propensity of different types of people for addictive behavior and the broader societal impact of their actions. Various solutions to the problem are possible, depending on whether addiction is viewed primarily as a matter of biochemical propensities, personality traits, material incentives, or socioeconomic structure. There will be both settled and contested facts concerning the implications of different policies. A shared representational framework constructed in the process of public deliberation can help render these aspects mutually intelligible, reduce the epistemic burden on behalf of policymakers, and reduce the psychological burden on behalf of citizens who can form their goals and aspirations without having to find different evaluative standards themselves. Furthermore, the participatory component of deliberation counters the normative challenge from above; it makes citizens design the policies, which increases policy acceptability and produces long-term compliance, as some studies suggest (Der Does and Jacquet, 2021; Fishkin et al., 2024).[12]

Conclusion

This article discussed autonomy as self-constitution, characterized by the competent mastery of the process of identity production through transpositional reasoning (Section Autonomy in BPP: Freedom of choice, agency, and self-constitution). We discussed epistemic, normative, and psychological challenges of this view of autonomy (Section The threefold challenge of autonomy as self-constitution). We then argued that the 'democratization of BPP' through deliberative mini-publics can help address, or at least mitigate, these challenges. In particular, they can assist individuals in reaching a *view from manywhere* by considering opinions and evaluative standards offered by other people, and investigating what kind of social environment they find conducive to their processes of identity formation.

Within the existing BPP toolkit, autonomy as self-constitution is most closely aligned with the ideas of boosting and nudge+. Boosts can help cultivate essential competencies for processes of self-constitution, and nudge+ highlights 'think' elements that complement nudging. Yet, both interventions primarily focus on *individual* reasoning and might underestimate the significance of the *social* dimensions of thinking and deliberating with others to uncover unexamined influences on people's preference formation processes. Our article highlighted that self-constitution extends beyond individual deliberation, occurring reflexively within broader social and institutional contexts. Overall, we have proposed to refocus – at least partly – the attention of BPP discussions from i-frame interventions to the question of s-frame interventions, i.e., questions of institutional design of public deliberation.

In doing so, our article introduces a twofold switch away from traditional BPP: from instrumental reasoning to substantial reasoning, and from interventions at the individual level to intervention in the institutional framework. If the goal of BPP is to enhance individual's autonomy, and autonomy requires not only to be able to have more opportunities but also being able to reason to become the author of one's own life, it requires the capacity to co-construct one's environment that will help one think about future preferences and goals. Therefore, we encourage questioning the traditional boundaries of BPP to effectively address autonomy, broadening the discussion to political economy and political philosophy.

We suggest emphasizing deliberation to democratize BPP. Admittedly, some individuals may not want to subject their process of identity formation to collective deliberation. Autonomy entails the freedom to resist collective decisions. Although this view has sometimes been used to negate the relevance of deliberation and public reasoning altogether (Sugden, 2018), we deem it important and sympathize with it. Autonomy involves the capacity to reason independently and avoid blindly obeying others or accepting their collective decisions without critical evaluation.

One concern that we have not sufficiently addressed in this context is the possibility that deliberation merely creates the illusion of autonomy. By involving citizens in deliberative forums, they may be led to believe they have genuine influence, while, in

reality, deliberation is used strategically by policymakers to secure public buy-in for decisions that reinforce the status quo. Dissenting voices can be neutralized by claiming that these individuals were 'heard' as part of the deliberative process, even if their input did not impact outcomes. While this risk of "repressive tolerance" (Marcuse, 1965) might be ameliorated by proper design principles which actively encourage progressive, emancipatory ideas, a proper focus on autonomy and deliberation in BPP calls for more work at the intersection of political economy, institutional analysis, and political philosophy. Such interdisciplinary work needs to assess how various institutional experiments can enhance a diversity of perspectives and enable pluralistic processes of identity construction while avoiding the danger of simply supporting a dominant status quo opinion or specifying a unique way of life others must follow.[13]

We would like to conclude with some final thoughts on why we believe that incorporating deliberation aligns with BPP's core idea of integrating insights from the behavioral sciences into public policy analysis. Unlike the rational choice framework, which assumes preferences are fixed, the concept of autonomy as self-constitution, as discussed in this article, broadens policy analysis to consider the formation of preferences and the factors that shape them. Behavioral scientists, particularly social psychologists, study how preferences and goals are formed (see, e.g., Sheldon, 2014; Dold et al. 2024). For instance, research in psychology suggests that non-conscious priming effects play a significant role in goal formation and selection (Bargh and Ferguson, 2000). Furthermore, goal formation and selection are deeply shaped by social and cultural contexts. Prevailing customs, norms, and institutions influence our cultural mental models—shared interpretive frameworks that encompass social identities, narratives, and worldviews—which, in turn, affect the way goals are formed and selected (Dold and Lewis, 2022). Importantly, we are largely unaware of both the goal-shaping power of these mental models and the cues that activate them in particular situations (Hoff and Stiglitz, 2016, p. 39). We argue that a BPP framework that values autonomy as self-constitution should prioritize raising awareness of individuals' cultural mental models and the subtle contextual factors that shape their goal formation. Standard BPP tools, such as nudges and boosts, can contribute to this awareness when complemented by deliberative elements. Our exploration of public deliberation as a crucial tool for fostering this awareness suggests the need to expand BPP, while still acknowledging and building upon the field's past achievements.

Data availability

No data were analyzed or generated in the writing of this article.

Received: 17 July 2024; Accepted: 4 March 2025;
Published online: 26 March 2025

Notes

1 For a summary of the defense, see Schmidt (2019); for a summary of the critique, see Rizzo and Whitman (2020).
2 There are exceptions, which directly endorse the paternalistic view, such as Conly (2013).
3 On the distinction between i-frame and s-frame policy interventions in the context of BPP, see Chater and Loewenstein (2023) and Connolly et al. (2024). Admittedly, many policy interventions have elements of both (Hallsworth 2023). Nevertheless, we think that the distinction is helpful to categorize different policy interventions along a continuum. In this article, we advocate for the field of BPP to give greater attention to policies that are closer to the s-frame end of the continuum.
4 Sunstein (2015, 2016) extends this approach, stating the importance of individualizing such interventions to realize personal autonomy rather than imposing a preconceived conception of the good on individuals.
5 In this article, we refrain from addressing the complex question of how to measure whether someone truly possesses autonomy, understood as self-constitution. However, we believe the conceptual and empirical work in Self-Determination Theory (SDT) offers valuable insights (Dold et al., 2024). SDT measures people's sense of autonomy by assessing the satisfaction of their basic psychological needs: autonomy, competence, and relatedness. Notably, individuals tend to score high on self-determination in institutional settings that provide basic civic freedoms and participatory rights (Ryan & DeHaan, 2023). This offers empirical support for the core argument of this paper, which emphasizes the link between public deliberation and individual autonomy.
6 See Dold (2024) for a deeper discussion of the value of education and experiments in living for autonomy as self-constitution.
7 See, for instance, Habermas (1996), Benhabib (1996), Gutman and Thompson (1997), Bohman (2004).
8 To address the epistemic challenge, recent evidence from the wellbeing literature highlights the effectiveness and practicality of participatory and deliberative policymaking (Fabian et al., 2022; Fabian et al., 2023). The authors advocate for a 'citizen perspective' (CP) over the traditional 'social planner perspective' (SPP) to better account for the value-laden nature of policy metrics. Their findings suggest that CP is not only feasible but also effective, with traditional skepticism about its practicality being largely overstated.
9 The capacity to learn from diversity is very important and a cornerstone of the epistemic democracy view, as defended by, Landemore (2020) or Goodin and Spiekerman (2018). The position often refers to the "diversity trumps ability theorem" derived from Hong and Page (2004).
10 That is the claim that it increases people's agentic capacity of transpositional thinking. On this point, see Dold (2024).
11 Mercier and Sperber (2011, 2017) emphasize that we engage in argumentative reasoning within groups, where one person presents their claim, and others evaluate it. When constructing an argument, our primary focus tends to be on persuasiveness rather than accuracy. Our main goal is to convince others, often leading us to selectively use evidence that supports our claim (a form of motivated reasoning). However, when we are the recipients of an argument, we aim to avoid being misled and thus practice "epistemic vigilance": we are on the lookout for signs of inaccuracy to reject the unconvincing arguments of others. This *asymmetry* benefits group reasoning and decision-making, particularly when diverse viewpoints are present.
12 If deliberative individuals choose to put forward routines or delegate the framing of the consumption of addictive substances to a third party, it is taking the idea of autonomy as self-constitution seriously. The decision not to reason before every potential encounter with addictive substances and to delegate is made deliberatively. In this sense, routinized choice and delegation need not be opposed to Autonomy III.
13 This concern can be mitigated in situations where exit is assured and a plurality of options is available. For a general defense of this view, see Taylor (2017).

References

Bächtiger A, Dryzek JS, Mansbridge J, Warren ME (2018) Deliberative democracy: An introduction. In: Bächtiger A, Dryzek JS, Warren ME, Mansbridge J (eds) The Oxford handbook of deliberative democracy. Oxford University Press, Oxford, pp 1–34
Banerjee S, Grüne-Yanoff T, John P, Moseley A (2024) It's time we put agency into behavioural public policy. Behav Public Policy
Banerjee S, John P (2021) Nudge plus: Incorporating reflection into behavioral public policy. Behav Public Policy 1–16
Bargh JA, Ferguson MJ (2000) Beyond behaviourism: On the automaticity of higher mental processes. Psychological Bulletin 126:925–945
Bartling B, Fehr E, Herz H (2014) The intrinsic value of decision rights. Econometrica 82(6):2005–2039
Benhabib S (1996) Democracy and difference. Princeton University Press, Princeton
Bermúdez JL (2020) Frame it again: New tools for rational decision-making. Cambridge University Press, Cambridge
Beshears J, Choi JJ, Laibson D, Madrian BC (2009) The importance of default options for retirement saving outcomes. Soc Secur Policy Chang Environ 167–195
Bobadilla-Suarez S, Sunstein CR, Sharot T (2017) The intrinsic value of choice: The propensity to under-delegate in the face of potential gains and losses. Journal Risk Uncertainty 54(3):187–202
Bohman J (2004) Realizing deliberative democracy as a mode of inquiry: Pragmatism, social facts, and normative theory. J Speculative Philos 18(1):23–43
Brennan J (2016) Against Democracy. Princeton University Press, Princeton
Chater N, Loewenstein G (2023) The i-frame and the s-frame: How focusing on individual-level solutions has led behavioral public policy astray. Behav Brain Sci 46
Chishima Y, Wilson AE (2021) Conversation with a future self: A letter-exchange exercise enhances student self-continuity, career planning, and academic thinking. Self Identity 20(5):646–651

Colin-Jaeger N, Dold M, Gascoin A (2022) Realism and deliberation in normative political economy: The fruitful intellectual dialogue between Rawls and Buchanan. In: Badei S, Griveaux A (eds) The positive and the normative in economics. Routledge, London, chap 6, pp 108–128
Conly S (2013) Against autonomy: Justifying coercive paternalism. Cambridge University Press, Cambridge
Connolly DJ, Loewenstein G, Chater N (2024) An s-frame agenda for behavioral public policy research. Behavioural Public Policy 1–21
Deci EL, Ryan RM (1987) The support of autonomy and the control of behavior. Journal Personality Social Psychology 53(6):1024–1037
Der Does R, Jacquet V (2021) Small-scale deliberation and mass democracy: A systematic review of the spillover effects of deliberative. Polit Stud 71(1):218–237
Dold M, Rizzo M (2021) The limits of opportunity-only: context-dependence and agency in behavioral welfare economics. J Econ Methodol 28(4):364–373
Dold M, Lewis P (2022) FA Hayek on the political economy of endogenous preferences: An historical overview and contemporary assessment. Journal Economic Behavior Organization 196:104–119
Dold M, Lewis P (2023) A neglected topos in behavioural normative economics: The opportunity and process aspect of freedom. Behav Public Policy 7(4):943–953
Dold M (2023) Individual Agency in Behavioral Public Policy: New Knowledge Problems. RG Working Paper. https://doi.org/10.13140/RG.2.2.12235.77603
Dold M (2024) Amartya Sen on Adaptation: Explications and Implications. RG Working Paper. https://doi.org/10.13140/RG.2.2.12081.60008
Dold M, van Emmerick E, Fabian M (2024) Taking psychology seriously: a self-determination theory perspective on Robert Sugden's opportunity criterion. Journal of Economic Methodology 1-18
Elster J (1986/2010) The market and the forum: Three varieties of political theory. In: The idea of the public sphere: A reader
Fabian M, Alexandrova A, Nair YC (2023) Coproducing Wellbeing Policy: A Theory of Thriving in Financial Hardship. Journal Happiness Studies 24(7):2309–2330
Fabian M, Alexandrova A, Coyle D, Agarwala M, Felici M (2022) Respecting the subject in wellbeing public policy: Beyond the social planner perspective. J Eur Public Policy
Fishkin J, Bolotnyv V, Lerner J, Siu A, Bradburn N (2024) Can deliberation have lasting effects? Am Polit Sci Rev
Gigerenzer G (2024) From bounded rationality to ecological rationality. In: Gigerenzer G, Mousavi S, Viale R (eds) The Herbert Simon Companion. Elgar Publishing
Goodin R, Spiekermann K (2018) An epistemic theory of democracy. Oxford University Press, Oxford
Grüne-Yanoff T (2021) Boosts: A remedy for Rizzo and Whitman's Panglossian fatalism. Rev Behav Econ 8(3–4):285–303
Grüne-Yanoff T, Hertwig R (2016) Nudge versus boost: How coherent are policy and theory? Minds Mach 26(1):149–183
Gutmann A, Thompson D (1997) Democracy and disagreement. Harvard University Press, Cambridge, MA
Habermas J (1996) Three normative models of democracy. In: Benhabib S (ed) Democracy and difference: Contesting the boundaries of the political. Princeton University Press, Princeton
Hallsworth M (2023) Behavioral public policy in practice: Misconceptions and opportunities. Behavioral & Brain Sciences 46
Hargreaves Heap SP (2023) Mill's constitution of liberty: An alternative behavioural policy framework. Behav Public Policy 7(4):933–942
Hausman DM, Welch B (2010) Debate: To nudge or not to nudge. J Polit Philos 18(1):123–136
Hertwig R, Grüne-Yanoff T (2017) Nudging and boosting: Steering or empowering good decisions. Perspect Psychol Sci 12(6):973–986
Hoff K, Stiglitz JE (2016) Striving for balance in economics: Towards a theory of the social determination of behaviour. Journal Economic Behaviour Organization 126:25–57
Hong L, Page S (2004) Groups of diverse problem solvers can outperform groups of high-ability problem solvers. Proc Natl Acad Sci USA 101(46):16385–16389
John P, Cotterill S, Moseley LRA, Smith G, Stoker G, Wales C (2011) Nudge, nudge, think, think: Experimenting with Ways to Change Civic Behaviour. Bloomsbury Academic
Kahneman D (2011) Thinking, fast and slow. Farrar, Straus and Giroux, New York
Knight J & Johnson J (2014) The priority of democracy: Political consequences of pragmatism. Princeton University Press, Princeton
Korsgaard C (2009) Self-constitution. Oxford University Press, Oxford
Kuyer P, Gordijn B (2023) Nudge in perspective: A systematic literature review on the ethical issues with nudging. Rationality Soc
Landemore H (2020) Open democracy: Reinventing popular rule for the twenty-first century. Princeton University Press, Princeton
Lanzing M (2019) "Strongly recommended" revisiting decisional privacy to judge hypernudging in self-tracking technologies. Philos Techno l 32:549–568
Lecouteux G, Mitroutchev I (2023) The view from manywhere: Normative economics with context-dependent preferences. Econ Philos 1:23
Lippmann W (1925) The phantom public. Routledge, London
Marcuse H (1965) Repressive tolerance. In: Wolff RP, Barrington MJ, Marcuse H (1970) A Critique of Pure Reason, Beacon Press, pp 81–123
McKenzie CR, Sher S, Leong LM, Müller-Trede J (2018) Constructed preferences, rationality, and choice architecture. Rev Behav Econ 5(3–4):337–360
Mercier H, Sperber D (2011) Why do humans reason? Arguments for an argumentative theory. Behav Brain Sci 34(2):57–74
Mercier H, Sperber D (2017) The enigma of reason. Harvard University Press, Cambridge, MA
Mill JS (1859/2003) On liberty. In: Bromwich D, Kateb G (eds) Yale University Press, New Haven and London
Müller, J. (2019) Political pluralism, disagreement and justice: the case for polycentric democracy. Routledge, London
Niemeyer S (2011) The emancipatory effect of deliberation: Empirical lessons from mini-publics. Polit Soc 39(1):103–140
Niemeyer S, Veri F, Dryzek J, Bächtiger A (2023) How deliberation happens: Enabling deliberative reason. Am Polit Sci Rev (Online version)
Noggle R (2018) Manipulation, salience, and nudges. Bioethics 32(3):164–170
O'Donoghue T, Rabin M (2006) Optimal sin taxes. J Public Econ 90(10–11):1825–1849
Rizzo M, Whitman G (2020) Escaping paternalism. Cambridge University Press, Cambridge
Ryan R, DeHaan C (2023) The social conditions for human flourishing: Economic and political influences on basic psychological needs. In: Ryan R (ed) The Oxford handbook of self-determination theory, pp 1149–1170. Oxford University Press
Saghai Y (2013) Salvaging the concept of nudge. J med ethics 39(8):487–493
Schauer F (1999) Talking as a decision procedure. In: Deliberative politics: Essays on democracy and disagreement. Oxford University Press, pp 17–27
Schmidt AT (2019) Getting real on rationality: Behavioral science, nudging, and public policy Ethics 129(4):511–543
Sen AK (2003) The importance of basic education. The Guardian
Sheldon KM (2014) Becoming oneself: The central role of self-concordant goal selection. Personality Social Psychology Review 18(4):349–365
Smith A (1859) The theory of moral sentiments. Clarendon Press, London
Sugden R (2018) Community of advantage. Oxford University Press, Oxford
Sunstein C (2002) On a danger of deliberative democracy. Daedalus 131(4):120–124
Sunstein C (2015) Choosing not to choose: Understanding the value of choice. Oxford University Press, New York
Sunstein C (2016) The ethics of influence: Government in the age of behavioral science. Cambridge University Press, New York
Sunstein C (2020) Behavioral welfare economics. J Benefit-Cost Anal 11(2):196–220
Taylor R (2017) Exit Left. Oxford University Press, Oxford
Thaler R, Sunstein C (2008) Nudges. Yale University Press, New Haven
Vugts A, Van den Hoven M, De Vet E, Verweij M (2020) How autonomy is understood in the discussion on the ethics of nudging. Behav Public Policy 4(1):108–123

Author contributions

Both authors, Nathanaël Colin-Jaeger and Malte Dold contributed equally to the work.

Competing interests

The authors declare no competing interests.

Ethical Approval

Ethical approval was not required as the study did not involve human participants

Informed Consent

This article does not contain any studies with human participants performed by any of the authors.

Additional information

Correspondence and requests for materials should be addressed to Malte Dold.