

# Can We Deliberate? How Motivated Reasoning Undermines Democratic Deliberation and What We Can Do About It



David R. H. Moscrop

Ever since democratic theory took the “deliberative turn” in the late 1980s and early 1990s, theories of deliberation have dominated discussions about democratic decision making (Dryzek, 2002; Goodin, 2008). Deliberative democracy principally concerns itself with generating or transforming individual and collective preferences through rational argument and the exchange of reasons by those affected by a decision (or by their representatives) (Elster, 1998; Fung, 2013; Gutmann & Thompson, 1996; Gutmann & Thompson, 2004; Rawls, 1996). Proponents of deliberative democracy justify their approach with one or more of three kinds of supporting argument: political, ethical, and epistemological. Each corresponds, respectively, to the expected ability of deliberative democracy to support and enhance democratic institutions, to generate fair and legitimate decisions that encourage compliance, and to produce epistemically “better” (i.e. more valid) decisions (Warren, 2002). This chapter is concerned with the third kind of supporting justification for deliberative democracy: the generation of epistemically valid judgments through the exchange of preferences and supporting reasons. While the epistemic argument for deliberation is normatively compelling, empirical research casts doubt on the extent to which the purported epistemic goods of democratic deliberation are being delivered across cases in practice.

Epistemically better judgments are valuable; if the claims made by proponents of the epistemic defense of deliberation are correct, properly-constituted deliberations will yield improved issue and preference understanding among those who deliberate (Chambers, 2006). If that is true, deliberation should produce better justifications for emergent preferences and more rational, acceptable, and legitimate policy options. However, if the claim is not true, or else if it holds less often in practice than expected, then proponents of the epistemic defense of deliberation will be required

---

D. R. H. Moscrop (✉)  
Department of Communication, University of Ottawa, Ottawa, ON, Canada  
e-mail: [dmoscrop@uottawa.ca](mailto:dmoscrop@uottawa.ca)

to revise the status of their claim on empirical grounds, work to find ways to bridge the gap between normative expectations and performance, or both.

In this chapter I ask *Can we deliberate?* I pose this question considering the specific challenge to deliberation by motivated reasoning—an affective, often irrational and non consciously directed, form of reasoning. This challenge is manifested in a threat to the autonomous agency of individuals required to *generate reasons that are recognized as reasons* that can be used to generate epistemically valid and authoritative judgments. Motivated reasoning is not the only troubling cognitive tendency that humans have—decades of research in social and political psychology have revealed many (see, for instance, Kahneman, 2011). But motivated reasoning is particularly troubling to the epistemic defense of democratic deliberation since as a common, stubborn cognitive process it may routinely undermine our ability, *outside of our own awareness*, to communicate openly and honestly with one another by undercutting what is meant to be an autonomous and rational process of intersubjectively establishing validity towards some particular end (for instance, generating a policy decision, establishing the rules of the game, or sorting out options to scrutinize). Motivated reasoning buries precisely what ought to be uncovered in democratic deliberation—our reasons and motivations for the preferences we develop and hold and the commitments that underwrite them.

To answer the question of whether we can deliberate autonomously considering motivated reasoning, I do four things. First, I argue in favor of a specific conception of personal autonomy that avoids two common sorts of definitional problems with the concept: overspecification and infinite regress. Second, I establish that epistemic deliberative democracy requires reasons generated by autonomous agents capable of connecting their motivations to their preferences and judgments in an intersubjective process of reason giving. Third, I explain how motivated reasoning undermines autonomous deliberation. Finally, I very briefly sketch three conceptual approaches to addressing the problems raised by the challenge of motivated reasoning to epistemic deliberative democracy as a first pass at reconceiving deliberation towards the end of more rational, autonomous deliberative practice.

## Varieties of Personal Autonomy

Part of the challenge of dealing with any discussion of autonomy is defining precisely what it means for someone to be autonomous. Autonomy is often loosely defined as a sort of freedom, liberty, or general absence of constraints—what I am calling *external* autonomy. For instance, someone is considered autonomous if she is free to cast a ballot in an election. In this chapter, I am not concerned with that sort of autonomy, which casts it as license rather than as a capacity. Rather, I am concerned with *internal* autonomy, which I define as having the ability to self govern one's thinking: the *capacity* to self-direct one's thoughts and actions *cognitively* and to justify those thoughts and actions to others with accurate reference to one's *true* reasons and motivations. So, one is *internally* autonomous to the extent that

they capable of and do form, direct, and control their cognition, even though throughout the cognitive process their thinking will interact be affected by factors including social groups, other individuals, physical locations, biological states, and so on. While internal autonomy may be best considered along a continuum—from more to less aware of what influences one’s cognition—it is probably a rare occurrence that someone is found *entirely* at one end of the spectrum or the other. Below, I will examine a definition of autonomy that is consistent with this approach. But first, we must consider two other approaches to characterizing autonomy.

Upon scratching the surface of what is required for internal autonomy, we find several justifiable approaches that do not fit together. Two common sorts problems plague definitions of autonomy: overspecification, in which the conditions of autonomy are so narrowly defined that internal autonomy is impossible, and infinite regress, the failure to specify an ultimate and decisive point at which the presence or absence of personal autonomy can be isolated and affirmed. One extreme conception of internal autonomy is known as “maximal autonomy” (Berofsky, 1995). This conception refers to radical, independent self-creation outside of any significant external direction or determination. This definition of autonomy, even as an ideal standard, suffers from the problem of overspecification. Indeed, the definition is implausible: socially, culturally, psychologically, and biologically it is an impossible standard to even approach, let alone meet, and it is not even clear whether it would be normatively desirable if we could. Moreover, it is unclear how one would measure compliance with attempts at pursuing it. Perhaps because of this major shortcoming, maximal definitions of autonomy are rarely used.

More commonly, autonomy is defined as a second-order capacity. But such capacity-based definitions tend to suffer from the problem infinite regress. For instance, Dworkin defines personal autonomy as “... a second-order capacity of persons to reflect critically upon their first order preferences, desires, wishes ... and the capacity to accept or attempt to change these in light of higher-order preferences and values” (Dworkin, 1988, p. 20). While this definition is more appealing and realistic than the maximal definition of autonomy, it still suffers from a logical deficiency: the problem of infinite regress. Because first-order preferences may be formed heteronomously, on a second-order level, any affirmation or rejection of a preference, desire, or wish will also require its own justification (and affirmation) if it is also to be an autonomous choice, and so on and on, ad infinitum (Christman, 1991). However, by slightly modifying this definition, we can address the problem of infinite regress.

Christman defines autonomy as a *process* that relies on a *capacity* rather than as a fixed state—and in this sense, it fits well with the continuum approach outlined above, since it is consistent with a dynamic and variable understanding of internal autonomy. It is a process of progressive checks that acts as a kind of cognitive reviewer and underwriter that draws on an individual’s ability to rationally self-assess. According to Christman, an individual is autonomous when “... the influences and conditions that give rise to the desire [or preference or intention] were factors that the agent approved of or did not resist, or would not have resisted had she attended to them, and that this judgment was or would have been made in a

*minimally rational, non-self-deceived manner*” (Christman, 1991, p. 22. Emphasis mine). Thus, in Christman’s formulation an agent can only be said to be autonomous if she is aware “of the changes and development of her character and of why they came about,” (Christman, 1991, p. 11) since only in this way can she encourage or resist these changes through deliberation. To present the argument as Christman does:

- (i) A person P is autonomous relative to some desire D if it is the case that P did not resist the development of D when attending to this process of development, or P *would not have* resisted that development had P attended to the process;
- (ii) The lack of resistance to the development of D did not take place (or would not have) under the influence of factors that inhibit self-reflection; and
- (iii) The self-reflection involved in condition (i) is (minimally) rational and involves no self-deception (Christman, 1991, p. 11. Emphasis in original).

His definition avoids the problem of infinite regress by fixing the conditions required for autonomy to the *first level* of evaluation: the *process* by which a desire, preference, or interest is formed (Christman, 1991, pp. 18–19). Thus, the process of evaluation, if undertaken in conditions of minimal rationality and self-awareness, serves as both the necessary and sufficient condition of autonomy without the need to evaluate any *particular* outcome.

Since so much in Christman’s argument depends upon precisely what is meant by autonomy, it is worth spending a bit more time specifying its definition, especially since autonomy is central to democratic decision making in general and deliberation in particular. To be more precise, autonomy requires what Christman calls “minimal ‘internal’ conditions for rationality” (Christman, 1991, p. 14). He cites a basic consistency of beliefs and desires as requirements but stops short of demanding that there be an absolute and clear link between the epistemic process of developing internal consistency and any presupposed ontologically objective account of the external world. Thus, autonomy requires only internal consistency and not a universal “objectively” verifiable connection to a pre-established shared reality.

The construction of autonomy I am working with in this chapter is consistent with Warren’s argument that deliberation is epistemically valuable to establish mutual understanding through talk (Warren, 2002). This is critically important. The link here to the model of deliberative democracy that I am working with is central to normative justifications for deliberation and we should expect to find it in practice: participants in deliberation who are in search of epistemic validity must be capable of maintaining at least a basic internal consistency that they can communicate with others who will recognize and adopt it; otherwise the ground upon which the deliberative enterprise rests dissolves. The decision, however, must reflect a logical consistency that is unlikely to emerge if it is drawn from a collection of inconsistent internal processes; and even if it did, it could not be said to be the product of an epistemically valid process.

Returning to Christman, for an agent to be autonomous “the influences and conditions” surrounding a judgment, through the interests, preferences, and desires that support such a judgment, must be approved of by the agent—or would have been approved of—under what we might call conditions of sufficient awareness (a

minimal level of knowledge about factors relevant to the judgment at hand). So, for an agent to be autonomous, she must be able to assent to all the immediate factors that contribute to a desire; proponents of democratic deliberation expect this capacity, since they expect that those who deliberate can and will give a true account of their reasons for or against a preference or proposition. The ability to deliver on these expectations requires, as Christman notes, a capacity for critical self-reflection (Christman, 1991, p. 11). To this point I must add a clarifying coda: for an agent to be autonomous she must also be able to access the factors that are *the actual mobilizing factors related to the desire or judgment at hand*. This requirement is centrally important to a robust defense of democratic deliberation. An agent is only autonomous to the extent that the process of self-reflection she undertakes in the process of approving of a desire or judgment *accurately* links “influences and conditions” to outcomes and is not interrupted or non-consciously distorted by some internal or external force. For instance, if a person tells you they support the Democrats or Republicans because of their capacity as economic managers or their position on school choice or abortion, but in reality supports them out of a deep, unspecified identity attachment, then that expression of support would not be autonomous (on party identity and partisanship see, for instance, Achens & Bartels, 2016). This requirement for autonomy is what I am referring to, drawing on Kant, as the principle of non-self-deception. It is premised on the hypothesis that while some degree of self-deception is a common enough occurrence among agents, it is possible to minimize how often it occurs and how significant its effects are when it does occur-again, to move the needle along the scale of internal autonomy close to the “autonomy” end and further away from the “heteronymy” end.<sup>1</sup>

## **Deliberative Democracy, Autonomy, and the Epistemic Defense of Deliberation**

Why is personal, internal autonomy important to and necessary for the epistemic defense deliberative democracy? Before addressing this question, it is necessary to define precisely what the epistemic defense deliberative democracy is. One of the leading and most comprehensive theories of epistemic deliberation is offered by Estlund (2008). In his conception of epistemic proceduralism, Estlund clearly and carefully distinguishes between purely procedural theories, which rely on some procedural good for legitimacy-such as Rawls’ “justice as fairness” approach (Rawls, 1999)-and thick, correctness-based epistemic theories-such as Rousseau’s General Will (Rousseau, 1971). While the former is only concerned with establishing fair procedures for decision making to generate legitimate outcomes, the latter requires

---

<sup>1</sup>For more on this, see Kant’s distinction between autarchy-the capacity to make decisions for oneself-and autonomy-the capacity to accurately give reasons for one’s decisions-in Elstub (2008), Guyer (2005), and Kant and Gregor (1785/1998).

that decisions be correct to be legitimate and authoritative: for instance, as the most popular product of a majority vote.

In contrast, Estlund's epistemic proceduralism requires only that the minority accept a majority decision as legitimate. They do not *have to* accept it as *correct* if it is the outcome a properly constituted and fair procedure *that has the tendency to generate correct outcomes* (within the parameters of a given, broader political or ethical system) *on a better-than-random basis*. The core of this argument rests of "the counterpart" of procedural fairness, epistemic proceduralism, which Estlund defines as: "procedural impartiality among individuals' opinions, but with a tendency to be correct; the impartial application of intelligence to the moral question at hand" (Estlund, 2008, p. 107). In the absence of a procedure-independent moral standard for producing and judging an outcome, this approach ensures that compliance based on the procedure's tendency to produce correct outcomes often enough, rather than mere fairness or the certainty that any decision produced must be morally correct—a requirement that necessitates some prior established standard that exists outside of deliberation (Estlund, 2008, p. 108).

So, why do theories of deliberative democracy, in so far as they are defended as epistemically superior approaches to generating valid and legitimate political decisions, require personal internal autonomy of the sort outline above? If the deliberative approach to democracy is to live up to its claim of producing correct decisions, even if on a mere better-than-random-chance basis, then that process will require that individuals connect their *actual reasons and motivations* to preferences that can be clearly communicated to others in a deliberative setting with fidelity to reality. That way, ultimately, the decisions that are generated through deliberation can be reasonably expected to link facts about the world—as they are interpreted and established by those assembled individuals—to reasons, then to their preferences, and, finally, to the decisions that are generated by the assembly. What autonomy ensures in this instance is the high-fidelity translation of the empirical and normative realities of deliberators into preferences, backed by motivations and reasons of which those deliberators are aware—which acts as a sort of first-level check on the validity of statements. What autonomy guards against is the presence of an internal (to the individual) fifth-column that acts to distort those empirical and normative realities, or that acts to mobilize them in a way contrary to the ultimate wishes of that individual, in such a way that the preferences generated by those individuals do not match their true preferences (i.e. those they would have chosen in a state of autonomy) and thus misleads the agent in question and subsequently all those with whom she engages on the issue.

In a deliberative assembly, the absence of a critical mass of autonomous individuals runs the risk that they outcomes they generate will be incorrect within the political or ethical parameters of the deliberation due to *structural distortion in judgment and decision-making tendencies*. This risk directly undermines the epistemic proceduralist defense of deliberation as an approach to decision-making that generates correct outcomes on a better-than-random basis. This is because the effects of the failure to reach a state of full autonomy (both in deliberative settings and non-deliberative settings) *are structural, rather than random*. The effects are *structural* in

two ways: first, they are structural in terms of *who they directly affect*; second, they are structural in terms of *who is affected by outcomes related to them*.

In the former instance, those most directly affected by a breakdown in autonomy—due to a lack of comprehension about the information they are using—tend to be less-educated, low information citizens (Althaus, 1998; Chong & Druckman, 2007; Converse, 1964; Cutler, 2002; Iyengar, Peters, & Kinder, 1982; Kuklinski & Quirk, 2000; Zaller, 1992). In the latter instance, those disproportionately affected by heteronomously generated preferences and outcomes tend to be people of color (Mendelberg, 2001; Snidennan, Hagen, Tetlock, & Brady, 1986), the poor and undereducated (Althaus, 1998; Zaller, 1992), and groups who tend to already suffer deleterious effects due to negative stereotyping (Kuklinski & Quirk, 2000). Thus, the issue of internal personal autonomy is normatively important in at least two ways: first, in the sense that it represents a commitment to a conceptual understanding of a good political life as being bound up in part in self determination; and second, in the sense that heteronymy might be bound up in the structural oppression, or at least the structural marginalization, of certain groups. While both normative approaches are interesting, it is primarily the former with which I am concerned in this chapter.

## Motivated Reasoning as a Challenge to Deliberative Democracy

### *What Is Motivated Reasoning?*

So far, I have argued that for theories of deliberative democracy to serve as plausible accounts of how to generate epistemically better judgments and decisions those who deliberate must be autonomous in a constrained sense. I have also claimed that the psychological phenomenon of motivated reasoning undermines autonomy and thus threatens to undermine the epistemic defense of deliberative democracy. In the following section I will explain what motivated reasoning is and outline specifically how it affects deliberators and undermines the epistemic authority of deliberations.

The phenomenon of motivated reasoning refers to “reliance on a biased set of cognitive processes: strategies for accessing, constructing, and evaluating beliefs” (Kunda, 1990, p. 480). Motivated reasoning serves as a core (potentially non-conscious) strategy employed by human beings in the interpretation of the world and the construction of reality—towards less accurate or more accurate, or better or worse, ends. In her review article on theory and evidence from the practice of deliberation, Tali Mendelberg (2002) notes, referencing Taber et al. (2001, p. 168), the bias in motivated reasoning “occurs at every step of information processing, from setting goals, to gathering and evaluating evidence from the outside or from memory, to constructing inferences and judgments.” Furthermore, Mendelberg cites Bodenhausen, Macrae, & Milne (1998, p. 169) to point out a few sources of



motivated reasoning: self-presentation, which comes from the desire to appear good, and self-deception, in which individuals are entirely unaware that they are in error or deluded. The latter is the sort with which I am concerned here.

There are two kinds of self-deceptive motivated reasoning: accuracy-driven and directional-driven. Accuracy-driven reasoning occurs when a subject is motivated to get the conclusion “right”—such as when there are rewards for a correct outcome or when the subject is required to justify their judgment publicly—and *tends to generate* more cognitively effortful, careful, and complex thinking, though there are limitations to avoiding biased cognition even under these circumstances (Kruglanski & Klar, 1987; Kunda, 1990). Think of a politician giving a speech in which his facts are being closely tracked and checked—he may be motivated to therefore get the facts right (although, these days, perhaps not). In the context of motivated reasoning, the word “bias” is, at least in its theoretical usage, stripped of any nonnative or ethically-evaluative (e.g. fair or unfair) content and instead refers to the phenomenon of prejudiced selection of tactics or strategies. Directionally-driven goals involve far less cognitive effort, care, and complexity; individuals who are directionally-driven are motivated to reach a defensible conclusion, but not necessarily the correct one. The last several years of American and European, Australian, Canadian, and other-politics are replete with examples, as any fact checking website reveals (see, for instance, [PolitiFact.com](http://PolitiFact.com) in the United States or [FactsCan.ca](http://FactsCan.ca) in Canada).

Under conditions of directionally-driven reasoning, individuals tend to maintain an “illusion of objectivity” (Pyszczynski & Greenberg, 1987) and reason in such a way that cognitive biases tend to be employed non-consciously—that is, without deliberate reflection and an active choice to employ them—and heavily directed towards maintaining or enhancing existing beliefs or preferences or reaching some desired conclusion (Kahan, 2013; Lodge & Taber, 2013; Redlawsk, 2002). While individuals reasoning towards directionally-driven goals cannot reach or justify *just any conclusion whatsoever*, they tend to reach use common tactics to reach more-or-less defensible, directionally-driven (and often biased) conclusions. Tactics for reaching such conclusions include selectively searching one’s memory for beliefs or evidence that support or confirm the desired conclusion, “creatively” combining and integrating existing and new evidence in such a way that supports the outcome they are after, and selectively choosing statistical heuristics that fit with their desired conclusions (Kunda, 1990, pp. 483–488). Challenging or disconfirming evidence in such cases tends to be ignored or rationalized in such a way that either minimizes conflicting or unwanted counter-evidence or else explains it away.

In summary, according to Kunda, “both kinds of goals affect reasoning by influencing the choice of beliefs and strategies applied to a given problem. But accuracy goals lead to the use of those beliefs and strategies that are considered most appropriate, whereas directional goals lead to the use of those that are considered most likely to yield the desired conclusion” (Kunda, 1990, p. 481). So, much of the reasoning we do is structurally biased; thus, when it comes to individual reasoning in a given political context, deliberative or otherwise, the question is not whether or not there is a biased—in the non-normative sense mentioned above—use of tactics or strategies for cognition, but whether the direction of cognition is towards accuracy or



some self-serving directional goal. We can now ask what the effects of motivated reasoning are on the practice of deliberative democracy and the kinds of judgments and decisions produced by democratic deliberation.

The sources of directional motivated reasoning vary, and each source poses different threats to epistemically good democratic deliberation. When it comes to political matters, attachment to a political party can lead to partisan motivated reasoning, which emerges from the relationship between an individual and their emotional attachment to a political party seen as representing or defending a group with which they identify (Leeper and Slothuus, 2014, pp. 136–137). Related to this challenge, individuals asked to evaluate candidates are cognitively biased in favor of their existing preferences, in stark violation of assumptions about the tendency of individuals to make fair, rational evaluations from memory rather than “hot” (i.e. affect-driven), immediate, and running evaluations with the goal of maintaining one’s existing preferences despite new or altered evidence (Redlawsk, 2002).

There are at least three structural cognitive errors, driven by motivated reasoning, occurring in these circumstances. The first is confirmation bias related to information search surrounding candidates they already approve of: individuals are looking to confirm what they already know and for which they hold positive affective evaluations (Redlawsk, 2002, p. 1025). The second, related to the first, is disconfirmation bias. In instances of disconfirmation bias, individuals will argue against, denigrate, explain away, or at least discount, information that challenges their pre existing preference and will actively seek to disconfirm evidence that violates their assumptions and preferences (Taber, Cann, & Kucsova, 2009, pp. 137–139). The third is the anchoring effect: voters, presented with negative evidence about a candidate they prefer, will *strengthen their support* for that candidate if they have already decided to support them (Redlawsk, 2002, pp. 1025–1026). Initial evaluations of a candidate, reinforced by biased searches for information and further biased processing of information about that candidate, create a strong anchor that either withstands challenges that should generate an adjust or, as noted, warps that adjustment so that evidence pushes an evaluation towards a candidate, when it should be pushing it away.

### ***Why Is Motivated Reasoning a Threat to Epistemically Good Democratic Deliberation?***

Epistemically good deliberation is deliberation that tends to bring about “correct” outcomes as defined here. The challenge of motivated reasoning to democratic deliberation is not one that can be explained away by asserting that the ideal of rational deliberation is a chimera, and that those who deliberate are necessarily human and thus predictably prone to a mix of rational and emotional thinking. Motivated reasoning is a specific and real challenge to deliberative democracy, not because it is emotional, but because *it is hidden*. The problem with motivated

reasoning is that it obscures motivations and makes it much more difficult for participants in a deliberation to put all their concerns and reasons on the table; under such conditions, it becomes difficult for each participant to access and engagement with the *actual* motivations and reasons that motivate a participant to hold this preference or to offer that reason. So, the question we should ask is not whether emotions or affective-based forms of reasoning—including motivated reasoning—can be eliminated from deliberation, but whether they can be interrogated, brought out into the open, and collectively managed.

We might assume that accuracy-directed cognition in a deliberative setting is epistemically desirable and that prompting accuracy goals is a solution to the challenge of motivated reasoning to epistemically good deliberation. After all, the epistemic concern of deliberation is getting to correct judgments and subsequent decisions. However, there are two problems with this. First, accuracy prompts do not always lead individuals to overcome the effects of motivated reasoning (Kunda, 1990; Redlawsk, 2002, pp. 1033–1035). So, the “cognitive baggage” that individuals bring with them into a deliberation cannot always be unpacked and set aside. And second, absent a consensus on what counts as “accurate” information within the group, the epistemic function of deliberation might be compromised and may result in a situation in which not only is a correct decision not reached, but sub-optimal outcomes are generated or exacerbated. In this case, a structural bias might be built into the structure of a democratic deliberation before it even begins and may influence proceedings from the get go. For example, a participant might deeply identify with a political party or organization and he may bring their views to the deliberation as his own, as a fixed heuristic. These two concerns encapsulate the general threat of motivated reasoning to deliberation: common, persistent or routine, and entrenched biased reasoning.

Motivated reasoning generates other threats to democratic deliberation. One threat related to the concern of the anchoring effect and generated by motivated reasoning is the “boomerang effect.” This occurs when some messaging strategy or approach inadvertently generates the *opposite* of the desired effect—and thus polarizes participants. So, in the context of a deliberation, some participants who are directionally-motivated may be induced to become further entrenched in their beliefs and less likely to support certain policies, even when presented with facts contrary to their existence beliefs or preferences or good arguments in favor of a certain policy—which can have a further polarizing effect (Hart & Nisbet, 2012). Indeed, in an experiment on motivated reasoning and preferences on climate change policy in the United States, Hart and Nisbet found that political partisanship influenced support for climate change and that new information—shared equally and presented identically—further *polarized* opinions on climate change between Republicans and Democrats (Hart & Nisbet, 2012; see also Taber, Cann, & Kucsova, 2009). If such effects are combined with a group in which a minority-type is outnumbered, polarization can become worse through increased (non-cognitive) bias and decreased cooperation (Bettencourt & Dorr, 1998).

Motivated reasoning also violates the basic deliberative requirement that strategic considerations be left outside the room, or, at least, that they are largely muted.

Instead, as noted, motivated reasoning threatens to introduce *structural* error into deliberation, threatening the epistemic force of deliberation as a theory of democratic decision making that produces correct outcomes on a better-than-chance basis. Regarding the epistemic function of deliberation, motivated reasoning undermines attempts at building shared understanding and generating shared preferences by increasing the chances of a boomerang effect occurring and by undermining the reliability of relatively stable, open-and therefore transparent-motivations that are open to discussion and debate. Moreover, motivated reasoning generates goals, incentives, and contributes to the generation of supporting arguments (or rationalizations) which may not be brought about absent such reasoning. Consequently, motivated reasoning *structurally violates* the principle of autonomy required for deliberation. When motivated reasoning is strong in a deliberation, the risk of a sort of shadow deliberation emerges: a kind of perverted deliberation that is taken place alongside the primary deliberation, based on factors and motivations that have little or nothing to do with the deliberation at hand. This is a significant threat to both the spirit and practice of deliberative democracy since if individuals were fully aware of the source of their motivations they might reason differently and present different preferences and justifications for those preferences.

## **How Should Theorists of Deliberation Respond to Motivated Reasoning?**

As challenging as motivated reasoning is to epistemic defenses of deliberative democracy, its existence alone is not enough to warrant dropping the epistemic defense of deliberation all together. For one, the challenge of motivated reasoning is probably *worse* in non-deliberative political settings. But, more importantly, it is likely that sophisticated deliberative design, both at the level of deliberative events (for instance, one-off deliberations or series of deliberations) and deliberative systems (for instance, institutionalized deliberative democracy, such as regular citizens' assemblies) can attenuate the deleterious effects of motivated reasoning, even if they are unable to eliminate them all together. If we believe that structure affects function-that the way something is designed will condition how it is used-then there may be several ways to change how cognition is "used" in deliberations, and thus to improve judgments made in democratic deliberation. Theorists of deliberation should respond to the challenge of motivated reasoning by changing how deliberations are structured and carried out. In this section I will briefly outline five approaches to minimizing the negative effects of motivated reasoning on the epistemic value of democratic deliberation. These approaches represent a mix of my own recommendations and those of others drawn from literature in social psychology and political science, combined in such a way as to directly address the challenge of cognitive distortion-in this case motivated reasoning-in the context of democratic deliberation.

## *Targeted Motivation*

As I have argued throughout this piece, inducing accuracy-driven goals is essential to moderating the effect of motivated reasoning, even if such inducements may not all together, or always, eliminate the challenge of such reasoning. Indeed, it is likely that it will require a combination of several approaches to seriously arrest the impacts of motivated reasoning on the production of epistemically good judgments in deliberative settings. Targeted motivation, however, is the first and most important approach to addressing this challenge. It relies on an understanding of the elaboration likelihood model (ELM), which was developed in the 1980s by psychologists Richard Petty and John Cacioppo. Petty and Cacioppo used dual-process theory in the model to specify two general routes through which a statement or argument might be processed: the central route, along which subjects were more likely to scrutinize a message, and a peripheral route along which subjects were more likely to employ cognitive short-cuts and external cues to evaluate it (Petty & Wegener, 1999; Petty, Cacioppo & Heesacker 1981; Petty & Cacioppo 1986). As the authors discovered, the key to getting subjects to employ the first route—one far better suited to the goals and exigencies deliberative democracy—was motivation: various factors, including a message's relevance or the availability of cognitive resources, went into determining which route a subject was likely to take (Chaiken & Trope, 1999; Jae & Delvecchio, 2004; Petty, Wells, & Brock, 1976).

Targeted motivation is an attempt to engage central-processing through highlighting to participants in a deliberation the relevance and importance of an issue and making it explicit that participants will be asked to explain and justify their preferences and underlying reasons to the gathered group. The goal of targeted motivation is to increase the probability that accuracy driven goals are primed prior to issue-related cognition. Targeted motivation can be further subdivided into tactics aimed at engaging individuals. Specifically, targeted motivation should take the form of ensuring that: (1) arguments are presented in clear, manageable form and language; (2) individuals are given appropriate amounts of time to scrutinize information, ask questions, and discuss their perspectives; (3) rewards for adopting peripheral methods are minimized or eliminated (e.g. rewards for finishing early or before another group or sub-group); (4) individuals are presented with clear arguments as to why a given issue is relevant to them, their families and friends, their community, city, state, or country; (5) the environment in which deliberation occurs is free from distracting elements, including any stimuli that may provide subtle nudges as to which way a participant should decide.

When it comes to motivated reasoning as a threat to autonomy and epistemically good deliberative judgments, within the context of deliberation, targeted motivation may assist in shifting subjects' attention towards the subject matter as well as *how they think about the subject matter*. This focus should help maximize the likelihood that individuals scrutinize the data and arguments presented to them, as well as bring some scrutiny to bear on their *cognitive process* for reaching a judgment. Targeted motivation alone may not entirely address the challenge of motivated

cognition by cueing accuracy-directed goals; it may, however, minimize instances of motivated reasoning. It may also enable other tactics, which I will discuss below, to work or else to work better.

### *Arational receptivity*

Motivated reasoning is driven by affect-and while much of the processing that occurs under conditions of motivated reasoning is hidden from consciousness, the *effects* of such are potentially traceable, if individuals are able and willing to interrogate them. Arational receptivity is a state in which individuals are open to publicly questioning and discussing-within the context of a group deliberation-their affective disposition towards issues and their related preferences. Cultivating openness to scrutinizing the affective dispositions one has towards certain issues and concomitant preferences could help generate stronger accuracy-driven goals and attenuate directionally-driven ones, especially in a public setting, by reminding participants and facilitators that when it comes to generating judgments affect is an entrenched and necessary element.

In fact, as neuroscientist Antonio Damasio has shown, feelings and emotions (what I lump together as “affect”) do important cognitive work and are essential for both mundane day-to-day choices as well as more complex decisions (Damasio, 1994, 2003). Just as notably, Heath and Pinker have neatly summarized, there are good evolutionary reasons why affect looms large in our lives: it is necessary for providing cues based on experience that are needed for future decisions (Heath, 2014; Pinker, 1997). This should not come as a surprise: brain systems drawing on affective considerations and processing tend to be faster and more efficient-if also more prone to error and bias-than those relying on rational reflection and processing (Kahneman, 2011). So, when it comes to the force of affect, we ought to row with the current, though we should also work hard to know where it is taking us. Returning to Kant: autonomy requires that individuals can both make decisions and have reasons for those decisions; I hasten to add that true autonomy requires that individuals have *valid and accurate* reasons for decisions, which would include the emotions and feelings that play a role in generating and perhaps sustaining those choices and which must be acknowledged if accuracy-driven goals are to be facilitated over directional-goals.

### *Cognitive Diversity*

Those who are subject to the effects of motivated reasoning and the directional-goals generated by it are, as I have discussed, also prone to polarization and boomerang effects under certain conditions when their reasons or preferences are challenged. However, as I have also mentioned, conditions generating

high-elaboration in cognition are more likely to generate accuracy-driven goals—which should mean that more autonomous deliberation would be brought about. So, a key challenge to overcoming the threat of directional motivated reasoning to the epistemic defense of democratic deliberation is finding ways to move individuals who deliberate from “cognitive auto pilot” to a more engaged and reflective state *without polarizing the group*. I believe that targeted motivation and arational receptivity are two important tools for this. But they may also require that groups be cognitively diverse.

According to Landmore (2013), who draws on Hong and Page (2004), cognitive diversity “refers to a diversity of ways of seeing the world, interpreting problems in it, and working out solutions to these problems. It denotes more specifically a diversity of perspectives ... interpretations ... heuristics ... and predictive models” (p. 1211). In deliberative contexts, the presence of cognitive diversity is hypothesized to improve decision-making (Page, 2008) and increase the quality of argumentation (Landmore, 2013). So, the presence of diverse ways of thinking may also offer a cognitive jolt to those who might otherwise rely heavily on the low-resource motivated reasoning when processing information and coming to judgments. Once again, the mechanism at work in such a case would likely be a shift from directional goals to accuracy goals; and the presence of a properly-constituted diverse group (Mendelberg, 2002) might enhance the effect of accuracy goals and further diminish motivated reasoning—potentially even eliminating, or at least significantly checking, the boomerang effect. At this stage, the cognitive diversity hypothesis is still largely experimental—it is indeed still a hypothesis—and more research is required into its long-term effects, plausibility, and generalizability. This is especially important in relation to how, if at all, cognitive diversity interacts with motivated reasoning and more specifically directional goals and polarization. None the less, if having diverse ways of approaching a problem means that individuals are more inclined to critically engage, and to consider closely their reasoning pattern, is promising.

## Conclusion

Motivated reasoning, insofar as it can be an irrational cognitive process, and because it occurs largely non-consciously, violates the principle normative goal of personal internal autonomy. This sort of autonomy is essential for generating epistemically good deliberative outcomes. Specifically, it does so by concealing the motivations of an individual who deliberates (i.e. to preserve and protect one’s current worldview) and biasing both the process of generating reasons for or against a preference and the reasons themselves. Of course, motivated reasoning is not universal and unavoidable; it can be attenuated. But as we have seen, it is commonplace, occurs in deliberations, and remains persistent in some cases despite efforts to counteract it with inducement to generating accuracy-driven outputs.

Occurrences of motivated reasoning act as challenges to democratic deliberation, but they do not render it useless—or even make it a less preferable alternative to aggregative democracy, which is easily worse at generating and exacerbating motivated reasoning in individuals engaged in political acts. Rather than gainsaying the value of such deliberation, the phenomenon of motivated reasoning points to an area of theories of deliberation—its epistemic defense—that requires further exploration and elaboration. It also serves as a reminder to both scholars and practitioners of deliberative democracy that deliberative design, at both the individual levels of deliberative events and the general level of institutional deliberative setup, that more work must be done if we are to generate the best possible outcomes from democratic deliberation.

## References

- Achens, L., & Bartels, C. (2016). *Democracy for realists*. Princeton, NJ: Princeton University Press.
- Althaus, S. (1998). Information effects and collective preferences. *American Political Science Review*, 92(3), 545–558.
- Berofsky, B. (1995). *Liberation from self: A theory of personal autonomy*. Cambridge, MA: Cambridge University Press.
- Bettencourt, B. A., & Dorr, N. (1998). Cooperative interaction and intergroup bias: Effects of numerical representation and cross-cut role assignment. *Personality and Social Psychology Bulletin*, 24(12), 1276–1293.
- Bodenhausen, G. V., Macrae, C. N., & Milne, A. B. (1998). Disregarding social stereotypes: Implications for memory, judgment, and behavior. In *Intentional forgetting: Interdisciplinary approaches* (pp. 349–368). Erlbaum.
- Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual-process theories in social psychology*. New York: Guilford Press.
- Chambers, S. (2006). *Reasonable democracy*. Ithaca: Cornell University Press.
- Chong, D., & Druckman, J. (2007). Framing public opinion in competitive democracies. *American Political Science Review*, 4(4), 637–655.
- Christman, J. (1991). Autonomy and personal history. *Canadian Journal of Philosophy*, 21(1), 1–24.
- Converse, P. (1964). The nature of belief systems in mass publics. In D. Apter (Ed.), *Ideology and discontent* (pp. 207–281). Glencoe, IL: Free Press of Glencoe.
- Cutler, F. (2002). The simplest shortcut of all: Sociodemographic characteristics and electoral choice. *Journal of Politics*, 64(2), 466–490.
- Damasio, A. R. (1994). *Descartes' error: Emotion, rationality and the human brain*. New York: Putnam.
- Damasio, A. R. (2003). *Looking for Spinoza: Joy, sorrow, and the feeling brain*. New York: Houghton Mifflin Harcourt.
- Dryzek, J. (2002). *Deliberative democracy and beyond*. New York, NY: Oxford University Press.
- Dworkin, G. (1988). *The theory and practice of autonomy*. Cambridge, MA: Cambridge University Press.
- Elster, J. (1998). Introduction. In J. Elster (Ed.), *Deliberative democracy* (pp. 1–18). Cambridge, MA: Cambridge University Press.
- Elstub, S. (2008). *Towards a deliberative and associational democracy*. Edinburgh: Edinburgh University Press.
- Estlund, D. (2008). *Democratic authority*. Princeton, NJ: Princeton University Press.



- Fung, A. (2013). The principle of affected interests: An interpretation and defense. In J. Nagel & R. M. Smith (Eds.), *Representation: Elections and beyond* (pp. 236–268). Philadelphia, PA: University of Pennsylvania Press.
- Goodin, R. (2008). *Innovating democracy*. New York, NY: Oxford University Press.
- Gutmann, A., & Thompson, D. (1996). *Democracy and disagreement*. Cambridge, MA: Belknap.
- Gutmann, A., & Thompson, D. (2004). *Why deliberative democracy?* Princeton, NJ: Princeton University Press.
- Guyer, P. (2005). *Kant's system of nature and freedom: Selected essays*. Oxford: Oxford University Press.
- Hart, P. S., & Nisbet, E. C. (2012). Boomerang effects in science communication: How motivated reasoning and identity cues amplify opinion polarization about climate mitigation policies. *Communication research*, 39(6), 701–723.
- Heath, J. (2014). *Enlightenment 2.0*. Toronto: Harper Collins.
- Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences*, 101(46), 16385–16389.
- Iyengar, S., Peters, M. D., & Kinder, D. R. (1982). Experimental demonstrations of the 'not-so-minimal' consequences of television news programs. *American Political Science Review*, 76(4), 848–858.
- Jae, H., & DelVecchio, D. (2004). Decision making by low-literacy consumers in the presence of point-of-purchase information. *Journal of Consumer Affairs*, 38(2), 342–354.
- Kahneman, Daniel. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.
- Kahan, D. M. (2013). Ideology, motivated reasoning, and cognitive reflection. *Judgment and Decision Making*, 8, 407–424.
- Kant, I., & Gregor, M. J. (1785/1998). *Groundwork of the metaphysics of morals*. Cambridge, MA: Cambridge University Press.
- Kruglanski, A. W., & Klar, Y. (1987). A view from the bridge: Synthesizing the consistency and attribution paradigms from a lay epistemic perspective. *European Journal of Social Psychology*, 17, 448–468.
- Kuklinski, J. H., & Quirk, P. J. (2000). Reconsidering the rational public: Cognition, heuristics, and mass opinion. *Elements of reason: Cognition, choice, and the bounds of rationality*, 153–82.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480–498.
- Landmore, H. (2013). *Democratic reason: Politics, collective intelligence, and the rule of the many*. Princeton: Princeton University Press.
- Leeper, T. J., & Slothuus, R. (2014). Political parties, motivated reasoning, and public opinion formation. *Political Psychology*, 35, 129–156.
- Lodge, M., & Taber, C. (2013). *The rationalizing voter*. Cambridge, MA: Cambridge University Press.
- Mendelberg, T. (2001). *The race card: Campaign strategy, implicit messages, and the norm of equality*. Princeton, NJ: Princeton University Press.
- Mendelberg, T. (2002). The deliberative citizen: Theory and evidence. *Political decision making, deliberation and participation*, 6(1), 151–193.
- Page, S. E. (2008). *The Difference: How the power of diversity creates better groups, firms, schools, and societies*. Princeton: Princeton University Press.
- Petty, R. E., Wells, G. L., & Brock, T. C. (1976). Distraction can enhance or reduce yielding to propaganda: Thought disruption versus effort justification. *Journal of Personality and social Psychology*, 34(5), 874.
- Petty, R. E., Cacioppo, J. T., & Heesacker, M. (1981). Effects of rhetorical questions on persuasion: A cognitive response analysis. *Journal of personality and social psychology*, 40(3), 432.
- Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In *Communication and persuasion* (pp. 1–24). New York: Springer.
- Petty, R. E., & Wegener, D. T. (1999). The elaboration likelihood model: Current status and controversies. *Dual-process theories in social psychology*, 1, 37–72.

- Pinker, S. (1997). *How the mind works*. New York: Norton.
- Pyszczynski, T., & Greenberg, J. (1987). Toward an integration of cognitive and motivation perspectives on social inferences: A biased hypothesis-testing model. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 297–340). New York, NY: Academic Press.
- Rawls, J. (1996). *Political liberalism*. New York: Columbia University Press.
- Rawls, J. (1999). *A theory of justice*. Cambridge, MA: Belknap Press of Harvard University.
- Redlawsk, D. P. (2002). Hot cognition or cool consideration? Testing the effects of motivated reasoning on political decision making. *The Journal of Politics*, 64(4), 1021–1044.
- Rousseau, J. J. (1971). In L. G. Crocker (Ed.), *The social contract, and discourse on the origin and foundation of inequality among mankind*. New York, NY: Pocket Books.
- Snidennan, P. M., Hagen, M. G., Tetlock, P. E., & Brady, H. E. (1986). Reasoning chains: Causal models of policy reasoning in mass publics. *British Journal of Political Science*, 16(4), 405–430.
- Taber, C. S., Cann, D., & Kucsova, S. (2009). The motivated processing of political arguments. *Political Behavior*, 31(2), 137–155.
- Taber, C. S., Lodge, M., & Glathar, J. (2001). The motivated construction of political judgments. In *Citizens and Politics: Perspectives from political psychology*, edited by J. H. Kuklinski. Cambridge: Cambridge University Press: 198–226.
- Warren, M. (2002). Deliberative democracy. In A. Carter & J. Stokes (Eds.), *Democratic theory today* (pp. 173–202). Cambridge, MA: Polity Press.
- Zaller, J. R. (1992). *The nature and origins of mass opinions*. Cambridge, MA: Cambridge University Press.