Agreeing to Disagree: Diversity, Political Contractualism, and the Open Society

John Thrasher, Chapman University

Political contractualism is important in societies characterized by substantial moral and political disagreement and diversity. The very disagreement that makes the social contract necessary, however, also makes agreement difficult. Call this the paradox of diversity, which is the result of a tension between two necessary conditions of political contractualism: existence and stability. The first involves showing the possibility of some agreement, while the second involves showing that the agreement can persist. To solve both of these problems, I develop a multilevel contract theory that I call the "open society" model of political contractualism that incorporates diversity into the contractual model at different levels solving the existence problem, while avoiding fragility in the face of the stability problem. This approach is able to take advantage of the benefits of institutional diversity while providing a stable framework for productive political disagreement.

D iversity—differences in fundamental religious, political, and moral beliefs and norms within a society—is at the heart of social contract theory and its political manifestation, liberalism. In diverse societies that do not share a uniform moral, political, or religious foundations some basis in agreement over stable rules and norms is necessary that does not appeal to the various doctrines on which we disagree. Political contractualism is the view that we need agreement on our basic political principles because we often agree on little else.[1] Those principles, constructed in the right way, allow us to disagree fruitfully and to organize ourselves cooperatively without the need to pursue a common social project over and above a commitment to the basic political principles that unify us.

Disagreement and diversity create the need for a social contract that can unify divergent interests and values into a common political system, but it is puzzling how people who fundamentally disagree about so much could come to any agreement at all. Call this the existence problem of political contractualism: is any agreement at all possible and, if so, under what circumstances? Much of the interesting work in contractualist theory is related to this existence question;

to devising models and experimenting with idealizations and specifications of the preferences, beliefs, and circumstances of the parties to the potential agreement to show that agreement is possible. A second question concerns the stability of any such agreement. Assuming an agreement is possible, will it survive in the face of change, both external and internal? Further, how robust is it in the face of diversity?

Given these two problems, there is a paradox at the heart of political contractualism in the face of diversity. Attempts to solve the existence problem often rely on assumptions that effectively reduce the diversity of the interests and beliefs of the parties to the contract. Even if this kind of strategy solves the existence problem, however, it can make it harder or impossible to solve the stability problem once diversity is reintroduced. I develop this problem before looking at how it is possible to solve it by embracing and harnessing diversity at several different theoretical and practical levels. I conclude by arguing that a social contract that takes diversity seriously will be the basis of a dynamic society not based on a fixed consensus but, rather, as a framework for an open society where we agree in order to disagree peacefully and fruitfully.

John Thrasher (johnthrasher23@gmail.com) is an assistant professor in the Philosophy Department and the Smith Institute for Political Economy and Philosophy at Chapman University in Orange, CA 92866.

1. It is common to distinguish between "contractarians" and "contractualists." The former are typically meant to be influenced by the Hobbesian interpretation of the social contract, for instance, James Buchanan (2000), David Gauthier (1986), and, more recently, Michael Moehler (2018) and Ryan Muldoon (2017); while contractualists have a more Kantian approach derived primarily from the work of T. M. Scanlon (1998). This division is unwieldy and misleading in any number of ways and I avoid it, calling all those who defend a contractual method of justification simply "contractualists." Here I follow Nic Southwood's (2010) ecumenism with regard to "contractualism."

THE PARADOX OF DIVERSITY

Political contractualism is a vision of a society of individuals cooperating together on mutually beneficial terms on which they can all voluntarily agree. It is a simple idea that has interesting implications. Abstracting from the particulars of many different contractualist theories, we can develop a general model of political contractualism that has the form:

> **General model of political contractualism.** N chooses R in M and this gives N^* reason to endorse and comply with R in the real world insofar as the reasons N has for choosing R in M can be shared by N^*.

This model is made up of two sets of agents (N and N^*), a deliberative model of agreement (M) and a set of target rules or principles (R). The agents in the deliberative model are model choosers that are meant to represent the reasons of actual people outside of the deliberative model (N^*). This model can take any number of forms based on its parameters and specifications, though the exact details of a particular version of contractualism are not important at this stage. This model is represented in figure 1.

Once there is a set of social rules (R) that individuals (N) can agree to, the existence problem has been effectively solved. But, since political contractualism is meant to be a normative standard for some existing or potential society, the model can only serve its function if it is relevantly related to the real people and situations it is meant to represent.[2] There must be a relationship between the real people looking at the standards of their own society and the individuals in the contractual model. Specifically, there needs to be a representation function between the reasoning of the individuals in the model agreement procedure and actual individuals. This relationship explicates the relationship between the reasons of the agents in the contractual model and the reasons of real people outside the model. The reasons of the representatives in the choice situation are meant to model our reasons; at least the reasons we have as political agents or citizens. As Rawls notes, contractual justification is not a "mere proof" (1999b, 508), nor is it merely reasoning correctly from given or generally accepted premises to conclusions (1980, 518). Rather, the normative task is to make explicit the reasoning that connects our standpoint as persons with de-
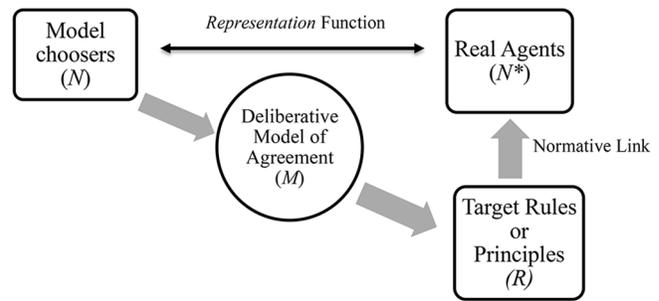

Figure 1. General model of contractualism

terminate interests and goals to our standpoint as citizens. The contractual model is a "device of representation" that connects our reasons with the reasons of our representatives in the contractual choice procedure (Rawls 1996, 24–27). This connection between the reasons of agents in the model and reasons of ordinary persons is the link that gives rational choice in the model normative power.[3]

To show that this kind of justification is possible, the political contractualist must show two things: (1) that some agreement on R in M is possible and (2) that the rules or principles agreed to (R) would continue to engender support in the face of diversity. Call the first the *existence problem* and the second the *stability problem*.

One way to solve the existence problem is to smooth out the differences between the parties to the agreement. This can be done in two ways. One is to impose limits on what the contractual agents can want or know, that is, reduce evaluative and/or doxastic diversity. Reducing doxastic diversity often has the effect of also limiting evaluative diversity and vice versa. The veil of ignorance, for instance, limits what contractual representatives know about themselves but also limits what they can plausibly value. People coming together to choose what rules they can live under will be unable to choose rules that benefit their own distinctive interests insofar as they do not know precisely what their interests are likely to be and precisely how the rules will affect those interests. As Rawls argues, reducing the information available to the parties is essential to generating a determinate result in the original position (1999b, 121).

This strategy to solve the existence problem, however, creates a tension with the stability problem. Reducing diversity may make the first problem more tractable, but given the normal functioning of a free society, diversity will inevitably reemerge. Rawls, recognizing this problem, developed

---

2. On this representation condition and the special problems of relating the reason of N to N^*, see Thrasher (2019). This conception of representation raises its own questions, but it is not totally distinct from other senses of representation in political theory (Rehfeld 2011, 2017) and in the philosophy of science in relation to "models" generally (e.g., Cartwright and Le Poidevin 1991; Colyvan 2013; Giere 2004; Weisberg 2007a, 2007b).

---

3. A core aspect of this representation condition is what Rawls calls the "publicity condition." Sam Freeman (2007) makes the case that publicity is essential for the normative link between model contractors and actual people.

a more diversity friendly version of political contractualism in his later work. Even there, however, Rawls (1996, 58) is clear that he is only concerned with what he calls "reasonable" pluralism. The modifier "reasonable" and it meaning has spawned a substantial literature in its own right and, despite this, there is still no consensus on how to precisely characterize this idea. Scanlon (1998) has, in his version of contractualism, done the most to make sense of what "reasonableness" amounts to, but in so doing he also creates a unique version of contractualism that poses its own theoretical challenges. One way to make sense of "reasonableness," or at least the role it plays in Rawls's theory, is to see it as a less radical version of the veil of ignorance, in the sense that it reduces the possible diversity of values and beliefs that contractors may have. The key difference is that reasonableness is not an informational constraint on what the contractors know or over the possible outcome they can rank; it does not unify their utility functions and their reasons for choice in the original position. Instead, reasonableness is a constraint on the possible utility functions of the choosers in the original position and in society generally. It functions as a way to reduce evaluative diversity in the model of choice. This restriction of diversity makes it possible, according to Rawls, for a society of people who may otherwise disagree about the good and much of the right to be able to find some core that can serve as a basis of an overlapping consensus that will remain stable over time.

So far, we have seen two main forms of diversity: doxastic and evaluative. Doxastic or informational diversity is reduced in contractual theories through some device like a veil of ignorance (Rawls 1999b) or veil of uncertainty (Brennan and Buchanan 2000; Buchanan 2000). Evaluative diversity is managed by constraining the utility functions or preferences that individuals may hold. We are now in a position to see how the strategy of reducing diversity in order to solve the existence problem is in tension with solving the stability problem. Call this the *paradox of diversity*. It is composed of several claims.

1. Diversity generates the conditions that make a social contract necessary.
2. Diversity creates disagreement, making it more difficult to find agreement on any social contract (the existence problem).
3. Solving the existence problem requires that one or both forms of diversity be reduced.
4. Stable agreements must engender their own support over time in the face of diversity.
5. Solving the existence problem by reducing diversity makes any agreement fragile in the face of diver-

sity, thereby making the stability problem harder to solve.

If one reduces diversity in order to generate agreement, not only is the point of generating agreement undermined, it may also make the contract less stable. Any agreement that requires limited diversity will likely be fragile in the face of increased diversity. On the other hand, if one does not reduce diversity, agreement may be impossible or, at least, it might be impossible to show that any agreement will be possible or unique.[4]

This is certainly a core problem in contractualist theory, but it is more than that; it is also an important practical political problem. Diversity can creep into the political system in any number of ways. Even a society that is relatively homogeneous will tend to drift into more diversity over time. Things tend toward diversity and, in a thermodynamic sense at least, disorder. This is welcomed, with perhaps only two cheers, as the inevitable "creative destruction" of a free society (Schumpeter 1994).[5] Freedom of speech, expression, and media make the creation and transmission of new ideas easy and natural. Even if somehow this process could be checked, perhaps by an extremely authoritarian state, diversity would likely emerge from external influences—through trade or other interactions with outside societies (Sparrow 2016).

This paradox does not claim that solving the existence problem will necessarily undermine stability. Instead, it shows that the tension between stability and existence is a potentially serious and perennial political problem—one with important implications. The key point is to recognize that existence and stability are two necessary conditions of a successful political contractualist theory that are linked together and related to diversity. We cannot simply solve the existence problem without considering the stability problem or vice versa.

The typical strategy to approaching these two problems has been to try to solve the existence problem and then show how that solution can be stable under various conditions. In the next several sections, recognizing the relationship between existence, stability, and diversity, I will argue that we can take a different approach and look to solve the existence problem by first looking at stability. This is because stability

---

4. The question for uniqueness in the solution to the existence problem may be creating many of the problems that are raised in the paradox of diversity; see Thrasher (2014).

5. The ability to generate and thrive in the face of "creative destruction" is one of the key elements of what Douglass North, John Wallis, and Barry Weingast (2009, 116–17) call "open-access orders" and what I will describe, more generally, as the "open society."

in a free society has a peculiar form that requires diversity and change to be built into it. By recognizing this, the existence problem becomes more tractable, though different, once we approach it from the point of view of stability.

THE STABILITY PROBLEM

One solution to the problem posed in the last section is to deny that stability is an important feature of a well-functioning and just society. Good societies, or at least good places to live, tend to be those that are vibrant and changing, not those that stand still—too much stability can be a bad thing.

Rawls distinguishes between two kinds of stability (1996, 142), one, necessary and proper for a contractual society and the other, less so. The first is the practical problem of ensuring that the political principles or constitution can persist and "as long as the means of persuasion or enforcement can be found" to maintain support, the society is stable (Rawls 1996, 142). We can think of many nonliberal societies (e.g., DPRK, USSR, the PRC) that have been remarkably stable in this sense, despite not having the voluntary support of their people. Rawls contrasts this kind of stability with another, more normatively interesting, form of stability that he calls "stability for the right reasons" (1996, xliii). Stability in this sense is the ability of a social order to appeal to the free public reason of equal citizens (1996, 143). As Rawls (1995, 146) notes in his reply to Habermas, showing that stability for the right reasons is possible in a pluralistic, diverse society is crucial to the project of public justification. In this way, stability is essential for legitimacy in both the practical and normative sense. Stability is a test on the reasonableness of the "public moral constitution" of a society (Rawls 1974, 12; 1996, 143).

This normative conception of stability is not an idiosyncratic feature of Rawls's contractualism; it is shared by Gauthier (1986, 230) as well. For both Gauthier and, perhaps more surprisingly, Rawls, the concept of stability has its roots in the idea of "equilibrium" drawn from game theory and economics (Gaus and Thrasher 2015). As Rawls writes in *A Theory of Justice*, stability is an equilibrium where the principles of justice are "collectively rational," that is, rational from each person's point of view (1999b, 435). He directly compares his idea of the well-ordered society to the concept of equilibrium in economics (1999a, 235). Although this conception of stability is developed considerably beyond its use in economics, it may still be helpful to look at the way the idea is used by economists to identify various ways that states of affairs can be both stable and dynamically changing. There are many such stability concepts in economics and game theory, and the range of ideas should alert us to the possibilities that there may be many such concepts available.

One is the Nash concept of stability in noncooperative games. This idea of stability is an equilibrium concept whereby strategies are in equilibrium if no player will benefit from unilaterally changing strategies—if a strategy is the best reply to everyone else's best reply (Nash 1950). This is a core notion in economics, and it helps to define the properties of an economy in general equilibrium. This idea, however, is in an important sense static. It tells us that an economy, once it achieves a particular state will tend to stay there, but it does not tell us anything about the process that may or may not lead to that equilibrium state or how robust that equilibrium will be to internal or external changes. We can describe this Nash conception of stability as a fixed-point conception of stability. Whatever may change in the system or state of affairs, that system will be stable if there is some fixed point that does not change in the system. What Rawls calls "inherent stability" in *A Theory of Justice* is a good example of a fixed-point conception of stability (Thrasher and Vallier 2018).

Stability, however, need not be understood as a fixed point in the sense that the principles that govern a society must remain the same over time. This point is also instructive in the political context. The constitutional settlement in the United States, for instance, has not remained the same over the last several centuries, but, despite this, the United States has continued to be governed by the same general principles. The American example shows that stability need not be primarily concerned with remaining the same but, rather, with creating a constitutional framework under which the political system can evolve in a constructive way along with the needs of the populace.

For instance, one of the most important challenges facing contemporary democracies, especially in Europe, is the challenge of diversity posed by increased migration of people from other parts of the world and, importantly, other religious, moral, and political traditions. Recent migration has caused a huge increase in doxastic as well as evaluative diversity and threatens the stability of the existing European constitutional settlement. Evidence of this can be found in the increased popularity of radical populist movements on the left and right. The most salient recent examples include Marine LePen's Front National (FN) in France, Alternative für Deutschland (AFD) in Germany, Golden Dawn in Greece, and, arguably, the Corbynite Labour Party as well as the ruling faction of the Conservative Party in Britain. All of these have seen surprising electoral success in recent elections. The most surprising and, perhaps most troubling, reaction to increased diversity has been the election of Donald Trump in the United States.

To address the challenges that diversity poses, we can develop a notion of dynamic stability. Rather than an equilibrium

conception, dynamic stability should be understood as a robustness test on a political or social constitution.[6] Of course, dynamic stability in this sense will also assume a certain amount of disharmony in the society. This is true in federal constitutions where there is often some important tension between the different political levels (Levy 2007). It is not harmony then that defines the stability of a dynamic system. Instead, a constitution embodying a political contractual agreement is dynamically stable insofar as its existence conditions persist in the face of significant diversity. What "significant" amounts to here will vary, but a dynamically stable constitution should be robust in the face of the diversity it is likely to face at maximum levels of feasible freedom. Dynamic stability is the result of a relatively fixed constitutional or contractual structure that fixes, in a very general way, basic fundamental principles, while allowing the exact institutional structure of a society to change considerably within that structure. Open societies do not try to settle all important questions, even all the important questions of justice. Instead, they settle some essential questions and create a durable framework for answering, in a provisional way, other questions in the course of social and political life. Consensus, in this broad sense, is not the goal of an open society but, rather, ongoing, structured dispute.

Consensus is the gold standard of politics. To achieve consensus is to reach unanimity, which is to say universal assent. No policy or principles that are ratified by a consensus can claim to silence some voices, or to speak only for the majority. Or so it seems. In the search for consensus, we see the paradox of diversity raise its head once again. Consensus is difficult to achieve in the face of diverse values and beliefs (Dryzek and Niemeyer 2006; Gaus 2016b; Kogelmann 2017). Consensus must, in a diverse society, be bought at a significant price. As we saw with the paradox of diversity, disagreement and diversity can be reduced or eliminated to generate agreement or consensus, but the piper needs to be paid at some point—diversity will inevitably reenter the political system. When it does, political systems that have been painstakingly constructed to generate agreement or consensus given certain assumptions will likely not remain stable in the face of increased diversity. Consensus-based theories of stability will tend to be fragile (Thrasher and Vallier 2015).

_____

6. Ryan Muldoon (2017) develops an innovative dynamic social contract theory that, nevertheless, relies on what I am calling a static equilibrium concept, specifically the Nash bargaining solution. His theory, as a whole, is dynamic because the social contractors—or at least their perspectives—are constantly changing and, hence, the bargaining solution generates different results over time. His theory does not propose a "final" social contract that is meant to be stable in the face of changing perspectives (Muldoon 2017, 120).

THE EXISTENCE PROBLEM

Although we are often concerned with agreement at the level of the social contract in our theories, it is at the democratic, institutional level that the real benefits of diversity can be brought to bear. In a general sense, institutions can be understood as structures of norms or stable equilibrium strategies (Schotter 2008). Thinking about political contractualism in the context of institutions may show how diversity can be integrated into contractualist theory while avoiding the paradox of diversity. Much of political theory, especially in the realm of democratic theory and social choice theory, sees political outcomes as a function of the beliefs and preferences of citizens filtered through some aggregation rule. Here, we will look at things differently. Imagine political institutions as a function that generates social outcomes given several constraints and parameters. The input is still the beliefs and preferences of individuals, but the domain of options is constrained by the foundational contract and constitutional structure. The institutions will set the agendas and the aggregation rules, and filter the results. To evaluate any outcomes, then, will require referencing the institutions in question.

Without going beyond the scope of this paper to discuss the exact nature of these institutions, we can highlight three features that would contribute to stability by harnessing diversity to generate better political outcomes. These are innovation, accountability, and robustness.

INNOVATION

The first and most obvious benefit of diversity in the context of institutions is innovation. Only when there are genuine differences in opinion and ideas about important matters is it possible to generate novel ideas and different approaches to old problems. This idea is at the heart of the Millian defense of freedom and diversity. In the contemporary context, Elizabeth Anderson (2006), Fred D'Agostino (2010), Gerald Gaus (2016b), Hélène Landemore (2012), Ryan Muldoon (2017), and others have argued that a similar defense of an innovative political system has many virtues.

In addition, many have touted the ability of diversity to aid democratic decision making. These approaches can be lumped together under the name "diversity trumps ability" (Anderson 2006). The core idea comes from work by Scott Page and his collaborator Lu Hong (2004). Although their model is somewhat complex mathematically, the basic idea is that a group of problem solvers who see the problem in question differently (but not too differently) will tend to outperform and, in some special circumstances always outperform, more knowledgeable or talented but less diverse groups. These results have been adapted and applied in the

context of democratic decision making by several thinkers. Hélène Landemore (2012) uses the "diversity trumps ability" results of Hong and Page in defense of an inclusive form of diverse deliberative democracy. Elizabeth Anderson (2006) uses the results of Hong and Page to argue in favor of greater inclusiveness and integration of society as a whole. The benefits of diversity give us a powerful reason, she argues, for favoring institutions that bring different types of people together and for reducing various forms of social segregation (Anderson 2013).

Despite criticisms (Gaus 2016b; Thompson 2014), there is something important to the idea that diverse groups of people can solve some problems better than more homogeneous groups. The important question is how and at what margin. Even if the particular theorem from Hong and Page is not the right model for thinking about doxastic or evaluative diversity, there are other models that tout the value of diversity in other ways (Thoma 2015; Weisberg and Muldoon 2009; Zollman 2009). These, however, also have their critics (Bruner 2019; O'Connor 2017; O'Connor and Bruner 2019).

Regardless of the particular model, the basic idea is that diversity, in the right institutional environment, is essential to innovation. The benefits of innovation will not always be shared by all, and those with conservative temperaments will likely rebel against innovation in general. Still, as Mill and others point out, innovation is crucial to the future of our society since we do not yet know what future challenges we will face or what possibilities are available to us. There is no way to accurately calibrate the level of innovation to ensure that it only generates results that we want. The open society, in this sense, is a society that is open not only to new ideas but also to disagreement. Only a robust marketplace of ideas will be able to make the case for some and against other innovations.

ACCOUNTABILITY

This idea of maximum diversity and innovation can be coupled with the importance of humility and epistemic fallibalism. This, too, is a core Millian (and Popperian) idea. We improve as a society by seeing which of our ideas are wrong, outdated, or no longer useful. The "selection pressures" in the biological analogy occur as ideas are subjected to public scrutiny or, to echo Thomas Jefferson, to the "bar of public reason." There is considerable evidence that individuals will consider evidence more carefully and deliberate more when they are expected to justify their conclusions publicly (D'Agostino 2010, 65–66). Making people accountable for their ideas in public can provide the selection pressure necessary to filter and sort ideas.

This is important because of recent findings in moral and political psychology suggesting that reasoned deliberation about morals or politics may be the exception rather than the rule. In particular Jonathan Haidt and his collaborators have shown that especially when it comes to political or moral conclusions our judgments come first and reasoning comes afterward (Greene and Haidt 2002; Haidt 2001). That is, we make judgments intuitively or emotionally and then use reason to justify our conclusions to ourselves and others. According to Haidt, we reason like lawyers rather than judges, finding reasons for conclusions that we already believe. Josh Greene, using similar results, has argued that this is good evidence that much of our moral (and political) judgments are driven by mere affect and not by reason (Greene 2008). Jesse Prinz (2008) has gone even further and argued that all of morality is based solely on the emotions.

All of this looks like bad news for notions of accountability and liberalism insofar as the latter is linked to the project of subjecting political programs to the test of public justification and public reason. If what we call "public reasoning" is really just a clash of nonrational judgments and competing affects, any pretensions to the legitimacy of democratic governance on the basis of anything like public reason look shaky. This also seems to undermine the basic notions of accountability and transparency that underlie many democratic institutions as well. No matter how officially transparent our institutions are, decisions are still made in the murky recesses of our emotional minds.

This verdict is too quick, though, and the reason why highlights benefits of diversity and democracy, at least in the right kinds of institutional settings. Haidt and company may be right about the limitations of individual reasoners—it is too easy to convince ourselves that we are right. A similar case can be made against deliberative settings that aim at consensus. Cognitive limitations coupled with the problems Haidt poses often lead to "discourse failures" in deliberative democratic settings (Pincione and Tesón 2006) and even group polarization (Sunstein 2002). In these settings, diversity is also often disvalued and reduced (Landemore and Page 2014). Further, when information is costly, many of these problems will multiply and lead to worse (McCubbins and Rodriguez 2006) and more fragile group decisions.

There is a silver lining, however. We are very bad at finding reasons against our favored point of view, but we are very good at picking out the holes in other people's moral or political judgments. When we reason alone or when we reason in a group of like-minded fellows aiming at consensus, these features work against us. When, however, we can turn our sights on one another, we are able to find the weak spots in each other's arguments. This is the core idea behind

the "argumentative theory of reason" developed by Hugo Mercier and Dan Sperber (2011). According to Mercier and Sperber, human reason developed as a way to increase assurance about the fidelity of communication (2011, 60). We are able to pick apart potentially deceitful or less than accurate communication from others, and we are disciplined by the knowledge that others will be able to do the same to us. Haidt follows Mercier and Sperber in arguing that this feature of our reasoning can be useful in some social and political settings (2013).

If constructive institutional settings can be developed and refined where motivated groups can debate in front of the public, we can keep each other honest and motivate ourselves to look for better and more convincing arguments. As Mercier and Sperber argue, since people are skilled at both evaluating and producing arguments, "these skills are displayed most easily in argumentative settings" and "debates should be especially conducive to good reasoning performance" (2011, 62). Mercier and Sperber (2011, 71–72) show in a convincing tour of the literature that by reasoning together, or at least against each other, we reason better because we are able to hold each other accountable. This leads to a virtuous cycle, an arms race where we seek to develop better and better arguments and inquire into the weaknesses of our own arguments. This process neither requires the public to be reasonable nor guarantees that reasonableness will be the result. Instead, it is a mechanism for finding and exposing error. Even in its purest form in the institutions of science, the result is not always truth or consensus, but what makes this mechanism robust is that it always has the means of discovering its own failings and exposing them by relying on the incentives of those involved in the dispute.

We might say that it is not by the love of truth that we get better arguments, but by our desire to convince our fellows and to be held in high esteem.[7] Much work needs to be done on testing different institutional settings for the purposes of eliciting the best results along these lines, but the examples of open parliaments, congresses, and courts provide good starting points. By using what might seem like a private vice and turning it into a public virtue, the political institutions of an open society benefit from the innovation of maximal diversity, held accountable by this public process.

ROBUSTNESS

Diversity is a precondition of innovation and dynamism in a society. Accountability harnesses that dynamism into pro-

ductive channels. The final piece of the puzzle is how diverse and competing political ideas and values can be more robust than a less diverse mix. If it is more robust, this will give us reason to think that more diversity and innovation, harnessed in an accountable way, will make a diverse and dynamic society more secure and better able to deal with unknown challenges. We have already seen one way that diversity will make societies more robust in that it will encourage and allow innovation. Without knowing what the future will bring, only with constant innovation will the material be available to meet future unforeseen challenges.

In addition to innovation and disciplining arguments and ideas, diversity in institutions makes societies more robust by bringing a wider range of approaches to bear on a given topic or issue. This is the advantage that Haidt cites (2013, chap. 13) when he argues that having liberals, conservatives, and libertarians together is better than having any of these groups alone because each approach is lacking a something that the other has. Liberals are experts in seeking out injustice and harm but, unlike conservatives, do not value tradition or authority; neither group values freedom as much as libertarians. The others can supplement the blind spots in one political approach, and having each around makes the competing viewpoints better off, even though this is hard to see in the middle of a heated debate. Gaus (2016a, 282) agrees, arguing that a moral and political order composed of liberals and conservatives performs better—from a certain moral point of view—than a homogeneous order of either alone. Of course there are limits to the amount of diversity that will be beneficial (Gaus 2018; Gaus and Hankins 2016).

Robustness in a system can be defined as "the maintenance of system performance either when subjected to external, unpredictable perturbations, or when there is uncertainty about the values of internal design parameters" (Anderies, Janssen, and Ostrom 2004, 17). Robustness in this sense is slightly different from the traditional political notion of stability. A stable state may or may not be robust. Some stable states are fragile. We can measure stability, following the standards in biology and ecology by looking at how easy it is to transition from one stable state to another, "more stable systems having higher barriers to switching" (Ives and Carpenter 2007, 58). This is what is called Holling's resilience (Holling 1973).

This idea of resilience can be illustrated in the political context. Imagine three different potentially stable states in a political society S_1, S_2, S_3. The first is the status quo, the second is an alternative constitution, and the third is a state of complete disorder with no set constitution. We can imagine that each alternative is an equilibrium so that if the society is in one state, there is a tendency to remain there. Each is stable, but we can imagine that they have different stability properties

_____

7. On the importance of esteem, see Brennan and Pettit (2006).

in terms of Holling's resilience. For instance, it might be easier to shift from S_2 to either S_1 or S_3. Once the society is either in the first constitution or anarchy, it is harder to shift to S_2, though if society is in S_2 it is easy to shift to either the first constitution or anarchy. This situation could lead to oscillation from the second constitution to either constitution one or anarchy. Even though each state is somewhat stable, the transition thresholds between states are not uniform, and so resilience in this case is a complex function of the relationships between each state and threshold.

A system is robust when it can continue to function in a stable state in the face of some range of perturbations. Often, robustness is the result of redundancy in the system. If there is a failure in one area, an overlapping element can provide the same function. This redundancy can often seem inefficient if one only considers operation at peak effectiveness, though there may be overriding benefits (Anderies et al. 2004). A robust system will typically not perform as efficiently with respect to a chosen set of criteria as its nonrobust counterpart. However, the robust system's performance will not drop off as rapidly as its nonrobust counterpart when confronted with external disturbance or internal stresses. Political societies must be robust in the same way that a power grid, nuclear power plant, or airline navigational system should be. Local failures should be contained and redundant elements should fill in for nonfunctioning elements. Diversity can help this goal by creating overlapping functions and institutions within a society. When the political system is functioning well, these redundancies may seem messy or inefficient, but when the system is not functioning well, we can see their benefit. It is characteristic of social planners and theorists to underestimate the importance of robustness and, thereby, to also underestimate the importance of institutional diversity in the name of efficiency or elegance.

Another important aspect that generates robustness as well as innovation is federalism. In a federal structure where localities at various levels have distinct jurisdictions and administrative preemption, innovation, accountability, and robustness are amplified (Bednar 2005, 2009). Federal subdivisions can innovate in their policies in an attempt to reach better outcomes than their neighbors (Ostrom and Ostrom 1965; Ostrom, Tiebout, and Warren 1961). So long as they can capture many of the benefits of positive changes and be harmed by negative ones, localities can compete to attract residents and businesses. People can "vote with their feet" and move to more congenial areas (Ostrom et al. 1961; Tiebout 1956). This incorporates the innovative and inclusive elements of the decentralized system discussed above, while avoiding some of its problems with implementation. Localities also tend to be more transparent and responsive to their residents, increasing their accountability. By providing a "laboratory" for democracy in terms of innovation and by creating an escape valve for poorly functioning localities or experiments, federalism adds to the robustness of a diverse society (Levy 2007).

Diversity has certain important benefits, but accepting those benefits in institutional contexts means that we must reject the very narrow conception of stability as the property of a system to easily return to a particular equilibrium when perturbed, at least in the institutional context. This idea of stability is not compatible with a dynamic and diverse society focused on innovation. Stability needs to be dynamic or understood as the robustness of the basic constitution of the society to persist, while the institutions themselves may be constantly changing. Diversity disrupts patterns. The challenge of stability is to show that the existence condition and its corollaries can be maintained in the face of realistic diversity.

POLITICAL CONTRACTUALISM AND THE OPEN SOCIETY

At this point, it is worth remembering the paradox of diversity for political contractualism that started this investigation. Recall that the problem is the tension between the necessary contractual conditions of existence and stability. Reducing diversity makes agreement easier to come by but less stable. There are two general ways of dealing with this problem. The first is to try to harness or incorporate diversity into the contractual model either in the deliberative model (Binmore 1994, 1998; Muldoon 2017) or by showing the benefits of diversity in the postcontract institutions (Bruner 2014; Kukathas 2003; Landemore 2012; Landemore and Page 2014; Nozick 1974). The second is to eliminate diversity in some way that allows agreement but that does not undermine the reason for the contract in the first place or imperil stability (Gaus 2011; Moehler 2014, 2018; Southwood 2010). We saw specific versions of each of these options in the last two sections.

Rawls, especially in *Political Liberalism*, saw the force of this problem and tried to find a middle way. His later solution hinges heavily on the importance of the idea of "reasonableness" in partitioning and organizing diversity. This is especially true in the work of latter-day Rawlsian public reason theorists who increasingly rely on a shared consensus conception of public reason to filter out unreasonable views (Cohen 2009; Hadfield and Macedo 2011; Leland and van Wietmarschen 2012; Quong 2011). This approach, however, has the drawback of solving the paradox by reducing diversity, making the contract potentially fragile or unstable in the face of increased diversity. It also has no principled way

of specifying exactly where and how to draw the lines of "reasonableness."[8]

One potential problem with all of these approaches is that they attempt to solve the paradox of diversity by limiting or accepting diversity throughout the contractual system as a whole. This, however, is itself the root cause of the paradox since the existence problem is harder to solve at some levels of abstraction than at others. By breaking the existence problem into several distinct levels, however, we can perhaps solve the paradox of diversity without endangering stability. This solution involves changing the target of contractual agreement from what Rawls (1999b, 235) calls "the well-ordered society" with its fixed-point conception of stability to what I will call the "open society" with a dynamic conception of stability (Thrasher and Vallier 2018).[9]

We will return to the differences in conceptions of stability, but the important difference from the point of view of the existence problem is that the open society divides contractual labor into several parts, each with a different role and process. On this approach, diversity is relevant at some stages but not others. We can think of these three stages as relating to three levels of political organization: the foundational, the constitutional, and the institutional. The first involves a foundational limiting principle, while the second establishes basic rights and powers in a constitutional structure, while the third is concerned with directly implementing policy.

At the foundational level, the contractors decide the fundamental limiting principle(s) of the society. The limiting principle is a constraint on the constitutional system. The limiting principles allow for certain constitutional possibilities and exclude others. In an open society, this will mean, minimally, that individuals are understood as not having any natural authority over any other. This is the starting point of any form of contractualism. Openness as a property of the society as a whole is not itself a limiting principle. Instead, it is the result of a principle that does not allow any person or faction to limit access to political and social institutions on the basis of who they are, rather than what they

do or whom they can persuade. This limiting principle can be described as an impersonality condition (North, Wallis, and Weingast 2009). Impersonality protects each individual's ability to enter into agreements and to form organizations in the open society.

Agreement on the limiting principle at this level, however, is not an overlapping consensus in the Rawlsian sense. The contractors need not endorse the limiting principle from a shared conception of their individual accounts of the good. Instead, they may agree on the basis of unshared, though convergent, reasons.[10] Because of this, at the foundational level of the contractual model there is no reason to limit evaluative and informational diversity of the parties (N). Instead, the diversity of the target of agreement (R) is limited. Even with maximally diverse agents, agreement on some basic limiting principle should be possible. The most likely candidate for a limiting principle is impersonality. Since it makes no reference to any particular identity it does not exclude anyone in principle. We may think, however, that some contractors will have reasons to think that impersonality is illegitimate or is undermined by their other evaluative concerns. This principle, though, is extremely thin and is consistent with many different types of societies and ways of life. Perhaps it is true that some would rather live outside of society than endorse a basic limiting principle of this sort, but the practical as well as principled reasons for endorsing this limiting principle over no limiting principle seem overwhelming.

Two important differences between this procedure and Rawls's original position are that (1) contractors are not choosing complete conceptions of justice or regime types and (2) choice need not be made on the basis of shared reasons. The foundational limiting principle(s) of a society are much more limited than what Rawls's parties choose in the original position. These foundational principle(s) are crucially important, but theories of justice and regimes are underdetermined by any set of limiting foundational principles. As Rawls realized in his later work, many conceptions of justice are compatible, for example, with the idea of freedom and equality.

The second difference is that choice of the limiting principles need not be made on the basis of shared reasons. This second point was mentioned above, but it bears elaborating. In the open society model of political contractualism, each agent (N) in the contractual model (M) may endorse the fundamental principle(s) from a diverse point of view or perspective. So long as they converge on the same limiting

---

8. Compare, for instance, Scanlon's (1998, 2014) interpretation of "reasonableness" with Cohen's (2009) and Gaus's (1996, 2006). Each is sophisticated and plausible, but each version is, nevertheless, substantially different and liable to its own difficulties.

9. This specific conception of the open society is meant to be a competitor to the Rawlsian liberal conception of the well-ordered society. This distinction is made in detail in Thrasher and Vallier (2018). As the name suggests, this conception of the good society is inspired by the vision of a free and dynamic society defended by Karl Popper (2013) and more recently by Gerald Gaus (2016b). Neither, however, develop a precise vision of open society. The conception of the open society developed here is heavily influenced by the idea of an open-access order (North et al. 2009) and the properties that such an order will have in the face of diversity and change.

10. On "convergence" justification, see, e.g., D'Agostino (1996), Gaus and Van Schoelandt (2017), and Kogelmann and Stich (2016).

principle, it is irrelevant whether their reasoning is based on conventional reasoning (Hardin 2003; Sabl 2012), instrumental reasoning (Moehler 2014, 2018), reasoning from a moral point of view (Gauthier 2013b, 2013a; Harman 1975; Harsanyi 1982), or some combination (Gaus 2011). This is because the limiting principle is not itself a conception of justice but merely the foundational limits of a constitutional order. Consensus in the Rawlsian sense is not essential. So long as they see those principles as legitimate and recognize that they have reasons to endorse and comply with them, this is enough.

The next level of contractual agreement is "constitutional," where basic features of the political society are decided given the background of the limiting principles justified at the foundational level. The constitution in this sense specifies the basic political rights and powers of the political society. These are meant to apply over an indefinite number of situations and should therefore be very general. Indeed these rules should be what Hart (1961) called "secondary rules" for the political system rather than specific first-order regulations or laws. These rules fix the framework for the institutions and norms of real political and social life, the market, and social cooperation. The constitutional level allows for a more diverse approach when setting up the basic terms of the political and social constitution of the society.

An important point to note, however, is that both of these levels of contractual agreement are not ongoing. This has the benefit of avoiding the instability that the completely unbounded process of contracting might create. This approach, for instance, can be contrasted with Ryan Muldoon's (2017) model of an unconstrained, ongoing social contract. Of course, the secondary rules of the constitution will have rules for how the rules themselves will be changed, but a process for changing rules is not the same thing as an ongoing constitutional convention. In a sense the constitutional rules are always subject to change, but these changes are made as problems arise, not through a regular ongoing process.

Once the limiting principle and the constitutional structure have been selected, there is still the question of the exact nature of the institutions that determine the specific, primary rules that govern social life. Even within the constraints set at the foundational and constitutional levels, there is a wealth of different institutional possibilities, considerable indeterminacy, and room for variation. Institutions operate in a context determined by the constitutional structure, the beliefs and values of their citizens, and their historical context. The role of the theorists at this stage is to provide comparative institutional analysis and to evaluate the basic ideas of the constitution in the context of the beliefs and norms of the particular society.

The evaluative and informational diversity that was dampened in the constitutional level by the practical need for general, impersonal rules that can apply over an indefinite number of individuals, organizations, and cases reemerges at this stage. People as they are must be the material for actual institutions, and it is implausible to reduce the diversity of their viewpoints for the purposes of institutional comparison. The historical context will also be important in the creation of specific social rules. These institutions and the organizations that are formed in this context will create new situations that will often require constitutional changes. In this way, the institutional level will also influence the constitutional. The relationship between diversity and the different contractual levels in the open society is represented in figure 2.

Although this is only a sketch of the open society as a model of political contractualism, it has some important features worth highlighting. The first is that it is the political conception of the basic structure of society. The basic rights do not reference morality as an independent area of investigation or justification. They also do not require underlying moral disagreement to justify them as in Moehler's (2018) multilevel contract theory.[11] Second, the foundational limiting principle(s) generated at the initial level are meant to be perfectly general and invariant to additional diversity. This is the point of allowing maximal diversity in at the foundational level since the existence and stability problem in this way are united, at least at the foundational level. Third, this theory moves much of the importance of political contractualism from the highest, foundational, level where it is typically located in contemporary contractual theories to the constitutional and, importantly, the institutional level.

One question is whether the idea presented here is merely a variation of Rawls's "four-stage sequence" from *A Theory of Justice*. Although there is an obvious similarity in form, there are important differences. Rawls sees the four-stage sequence as a framework "to simplify the application of the two principles of justice" (1999b, 171) and as a more informationally rich "elaboration of the original position" (1999b, 172). Once the parties have selected the principles of justice in the original position, they move to their application in a

---

11. There are many differences between Moehler's multilevel contract theory and the one defended here. The first is that the open society is a form of political contractualism that makes no direct reference to morality or moral disagreement. Moehler requires moral disagreement at the lower level to generate the need to ascend to the instrumental level; the open society does not require or reference moral disagreement. Further, the principles generated in Moehler's contract (e.g., the weak principle of universalization and the stabilized Nash bargaining solution) go far beyond what the foundational level of the contract in the open society justifies.
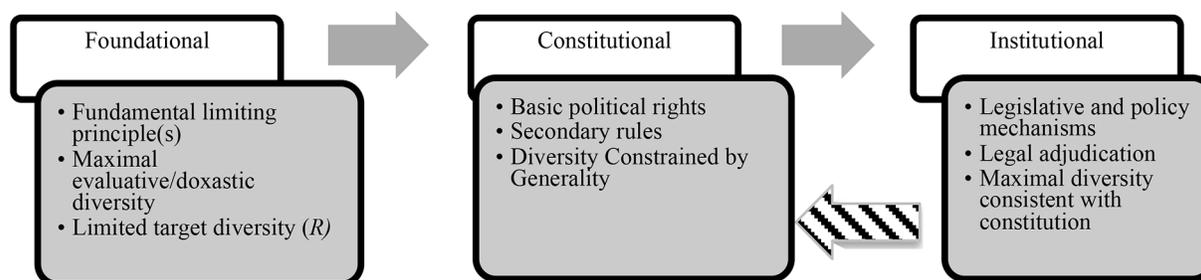
Figure 2. Diversity in the open society model

constitutional convention. The important point to note is that the principles are already chosen by the time of the constitutional selection; this is not true in the constitutional contract described above. The foundational contract only assigns the basic limiting principles of the society (e.g., impersonality). Unlike the first stage of the four-stage sequence, no determinate conception of justice is agreed to in the foundational contract.

Another important difference is that informational and evaluative diversity is not excluded in the foundational contract, unlike in the four-stage sequence. Since there is maximal diversity at this stage, contractors would limit disagreement in this stage by focusing on the highest-level limiting principle as a target of agreement without limiting diversity. Importantly, this reduction of disagreement is part of the process of contracting, not a modeling assumption about the values and knowledge of the contractors themselves—no veil of any level of thickness is required.

The final difference concerns Rawls's conception of the four-stage sequence and its relation to the principles of justice. In an important footnote, Rawls (1999b, 173 n. 3) distinguishes his approach in the four-stage sequence from the contractualism of Buchanan and Tullock in *The Calculus of Consent* (1999). As he puts it, the goal of the four-stage sequence is to assist in application of an already agreed upon conception of justice, "the aim is to characterize a just constitution," not to show that such a constitution would be agreed to "under more or less realistic (though simplified) assumptions about political life" (1999b, 173). This is a crucial difference between the contractual theory developed here and Rawls's four-stage sequence. It is exactly this sharp distinction between the existence and stability problems in Rawls (and other contract theories) that I have argued creates the paradox of diversity in the first place. The goal of the contractual theory presented here is to incorporate questions of stability into the existence question, thereby fusing the two types of questions that Rawls seeks to keep distinct.

It might seem that by incorporating concerns about stability into the existence question, the normativity of resulting

contractual agreement will be undermined. On the contrary, the normative force of the contractual agreement depends on how closely the reasons of the representative contractors are to the reasons of actual people to whom the contract theory is meant to apply. The link between these two sets of reasons is the normative link that allows this approach to serve as a justification for political institutions (see fig. 1). Making the tests for stability and existence harder by fusing them and allowing maximal diversity at the outset assures, in a way that other contractual theories cannot, that this approach is both maximally inclusive and robust. As a political standard for a liberal society, inclusivity is certainly a virtue, especially in modern pluralistic societies like our own. Robustness is assured since it is unlikely that some additional diversity will arise that was not already captured in the initial contractual situation. Both of these features should also enhance the stability of the contract.

It is the fusing of the existence and stability problems in the various levels of this approach that allows the open society to avoid the paradox of diversity that other forms of contractualism are susceptible to. At the foundational level, these questions are answered together, while at the constitutional and institutional level, the openness that is protected by the limiting principle allows the three benefits of openness (innovation, accountability, and robustness) to generate and reinforce their own support over time. The chief difference between the open society and contractual theories of liberalism or justice is precisely in the fact that no fixed-point conception of justice is agreed to in the open society's model of political contractualism. The open society, in this way avoids the fragility of fixed-point conceptions of justice as well as the instability inherent in societies whose basic terms are always in question.

DIVERSITY AND POLITICAL CONTRACTUALISM

I have argued that diversity creates an essential tension at the heart of social contract theory embodied in what I have called the paradox of diversity. To solve this problem, I proposed a form of multilevel political contractualism that

I have called "the open society." This approach solves the paradox by incorporating diversity into the social contract at different degrees at different levels, solving the existence problem while preserving stability in a dynamic, robust way. In addition, the focus on the institutional benefits of diversity in democratic contractualism point in the direction of how real institutions can harness diversity.

The open society model, if it is as viable as I have argued, shows that there are great advantages to living in a society characterized by significant and deep disagreement, if society is organized to take advantage of those benefits. As I have argued, it is essential that this disagreement and diversity take place in a constitutional structure that secures basic rights and creates a stable setting for dynamic institutions. In such a setting, disagreement and diversity are essential because they correct the blind spots we each have and introduce new ideas that help us see further than we might have on our own. It turns out that we need each other more than we think, not despite our disagreements, but because of them.

ACKNOWLEDGMENTS

REFERENCES

Anderies, John M., Marco A. Janssen, and Elinor Ostrom. 2004. "A Framework to Analyze the Robustness of Social-Ecological Systems from an Institutional Perspective." *Ecology and Society* 9 (1), article 18.

Anderson, Elizabeth. 2006. "The Epistemology of Democracy." *Episteme* 3 (1–2): 8–22.

Anderson, Elizabeth. 2013. *The Imperative of Integration*. Reprint ed. Princeton, NJ: Princeton University Press.

Bednar, Jenna. 2005. "Federalism as a Public Good." *Constitutional Political Economy* 16 (2): 189–205.

Bednar, Jenna. 2009. *The Robust Federation*. Cambridge: Cambridge University Press.

Binmore, Ken. 1994. *Game Theory and the Social Contract*, vol. 1, *Playing Fair*. Cambridge, MA: MIT Press.

Binmore, Ken. 1998. *Game Theory and the Social Contract*, vol. 2, *Just Playing*. Cambridge, MA: MIT Press.

Brennan, Geoffrey, and James Buchanan. 2000. *The Reason of Rules*, vol. 10. The Collected Works of James M. Buchanan. Indianapolis: Liberty Fund.

Brennan, Geoffrey, and Philip Pettit. 2006. *The Economy of Esteem: An Essay on Civil and Political Society*. Oxford: Oxford University Press.

Bruner, Justin P. 2014. "Diversity, Tolerance, and the Social Contract." *Politics, Philosophy and Economics* 14 (4): 1–20.

Bruner, Justin P. 2019. "Minority (Dis)Advantage in Population Games." *Synthese* 196 (1): 413–27.

Buchanan, James. 2000. *The Limits of Liberty: Between Anarchy and Leviathan*. The Collected Works of James M. Buchanan. Indianapolis: Liberty Fund.

Buchanan, James, and Gordon Tullock. 1999. *The Calculus of Consent: Logical Foundations of Constitutional Democracy*. The Collected Works of James M. Buchanan. Indianapolis: Liberty Fund.

Cartwright, Nancy, and Robin Le Poidevin. 1991. "Fables and Models." *Proceedings of the Aristotelian Society, Supplementary Volumes* 65:55–82.

Cohen, Joshua. 2009. "Truth and Public Reason." *Philosophy and Public Affairs* 37 (1): 2–42.

Colyvan, Mark. 2013. "Idealisations in Normative Models." *Synthese* 190 (8): 1337–50.

D'Agostino, Fred. 1996. *Free Public Reason: Making It Up As We Go*. Oxford: Oxford University Press.

D'Agostino, Fred. 2010. *Naturalizing Epistemology: Thomas Kuhn and the "Essential Tension."* London: Palgrave Macmillan.

Dryzek, John, and Simon Niemeyer. 2006. "Reconciling Pluralism and Consensus as Political Ideals." *American Journal of Political Science* 50 (3): 634–49.

Freeman, Samuel. 2007. "The Burdens of Public Justification: Constructivism, Contractualism, and Publicity." *Politics, Philosophy and Economics* 6 (1): 5–43.

Gaus, Gerald. 1996. *Justificatory Liberalism: An Essay on Epistemology and Political Theory*. Oxford: Oxford University Press.

Gaus, Gerald. 2006. "The Rational, the Reasonable and Justification." *Journal of Political Philosophy* 3 (3): 234–58.

Gaus, Gerald. 2011. *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World*. New York: Cambridge University Press.

Gaus, Gerald. 2016a. "The Role of Conservatism in Securing and Maintaining Just Moral Constitutions: Toward a Theory of Complex Normative Systems." In Sanford Levinson, Parker Joel, and Melissa Williams, eds., *Nomos LVI: American Conservatism*. New York: New York University Press, 256–91.

Gaus, Gerald. 2016b. *The Tyranny of the Ideal: Justice in a Diverse Society*. Princeton, NJ: Princeton University Press.

Gaus, Gerald. 2018. "Self-Organizing Moral Systems: Beyond Social Contract Theory." *Politics, Philosophy and Economics* 17 (2): 119–47. https://doi.org/10.1177/1470594X17719425.

Gaus, Gerald, and Keith Hankins. 2016. "Searching for the Ideal: The Fundamental Diversity Dilemma." In Kevin Vallier and Michael Weber, eds., *Political Utopias: Promise or Peril?* Oxford: Oxford University Press.

Gaus, Gerald, and John Thrasher. 2015. "Rational Choice in the Original Position: The (Many) Models of Rawls and Harsanyi." In Timothy Hinton, ed., *The Cambridge Companion to the Original Position*. Cambridge: Cambridge University Press.

Gaus, Gerald, and Chad Van Schoelandt. 2017. "Consensus on What? Convergence for What? Four Models of Political Liberalism." *Ethics* 128 (1): 145–72.

Gauthier, David. 1986. *Morals by Agreement*. Oxford: Clarendon.

Gauthier, David. 2013a. "Achieving Pareto-Optimality: Invisible Hands, Social Contracts, and Rational Deliberation." *Rationality, Markets and Morals* 4:191–204.

Gauthier, David. 2013b. "Twenty-Five On." *Ethics* 123 (4): 601–24. https://doi.org/10.1086/670246.

Giere, Ronald. 2004. "How Models Are Used to Represent Reality." *Philosophy of Science* 71 (5): 742–52.

Greene, Joshua. 2008. "The Secret Joke of Kant's Soul." In Walter Sinnott-Armstrong, ed., *Moral Psychology*. The Neuroscience of Morality: Emotion, Brain Disorders, and Development. Cambridge, MA: MIT Press, 35–80.

Greene, Joshua, and Jonathan Haidt. 2002. "How (and Where) Does Moral Judgment Work?" *Trends in Cognitive Sciences* 6 (12): 517–23.

Hadfield, Gillian, and Stephen Macedo. 2011. "Rational Reasonableness: Toward a Positive Theory of Public Reason." University of Southern California Law and Economics Working paper series, March. http://law.bepress.com/usclwps-lewps/art127.

Haidt, Jonathan. 2001. "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment." *Psychological Review* 108 (4): 814–34.

Haidt, Jonathan. 2013. *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. New York: Vintage.

Hardin, Russell. 2003. *Liberalism, Constitutionalism, and Democracy*. Oxford: Oxford University Press.

Harman, Gilbert. 1975. "Moral Relativism Defended." *Philosophical Review* 84 (1): 3–22.

Harsanyi, John. 1982. "Morality and the Theory of Rational Behavior." In Amartya Sen and Bernard Williams, eds., *Utilitarianism and Beyond*. Cambridge: Cambridge University Press, 39–62.

Hart, H. L. A. 1961. *The Concept of Law*. Oxford: Clarendon.

Holling, C. S. 1973. "Resilience and Stability of Ecological Systems." *Annual Review of Ecology and Systematics* 4 (1): 1–23. https://doi.org/10.1146/annurev.es.04.110173.000245.

Hong, Lu, and Scott Page. 2004. "Groups of Diverse Problem Solvers Can Outperform Groups of High-Ability Problem Solvers." *Proceedings of the National Academy of Sciences of the United States of America* 101 (46): 16385–89.

Ives, Anthony R., and Stephen R. Carpenter. 2007. "Stability and Diversity of Ecosystems." *Science* 317 (5834): 58–62. https://doi.org/10.1126/science.1133258.

Kogelmann, Brian. 2017. "Justice, Diversity, and the Well-Ordered Society." *Philosophical Quarterly* 67 (269): 663–84.

Kogelmann, Brian, and Stephen Stich. 2016. "When Public Reason Fails Us: Convergence Discourse as Blood Oath." *American Political Science Review* 110 (4): 717–30.

Kukathas, Chandran. 2003. *The Liberal Archipelago*. Oxford: Oxford University Press.

Landemore, Hélène. 2012. *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many*. Princeton, NJ: Princeton University Press.

Landemore, Hélène, and Scott Page. 2014. "Deliberation and Disagreement: Problem Solving, Prediction, and Positive Dissensus." *Politics, Philosophy and Economics* 14 (3): 229–54.

Leland, R. J., and Han van Wietmarschen. 2012. "Reasonableness, Intellectual Modesty, and Reciprocity in Political Justification." *Ethics* 122 (4): 721–47.

Levy, Jacob T. 2007. "Federalism, Liberalism, and the Separation of Loyalties." *American Political Science Review* 101 (3): 459–77. https://doi.org/10.1017/S0003055407070268.

McCubbins, Mathew, and Daniel Rodriguez. 2006. "When Does Deliberating Improve Decisionmaking?" *Journal of Contemporary Legal Issues* 15 (January): 9–50.

Mercier, Hugo, and Dan Sperber. 2011. "Why Do Humans Reason? Arguments for an Argumentative Theory." *Behavioral and Brain Sciences* 34 (2): 57–74.

Moehler, Michael. 2014. "The Scope of Instrumental Morality." *Philosophical Studies* 167 (2): 435–51.

Moehler, Michael. 2018. *Minimal Morality: A Multilevel Social Contract Theory*. Oxford: Oxford University Press.

Muldoon, Ryan. 2017. *Social Contract Theory for a Diverse World: Beyond Tolerance*. London: Routledge.

Nash, John. 1950. "Equilibrium Points in N-Person Games." *Proceedings of the National Academy of Sciences of the United States of America* 36 (1): 48–49.

North, Douglass C., John Joseph Wallis, and Barry R. Weingast. 2009. *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History*. 1st ed. Cambridge: Cambridge University Press.

Nozick, Robert. 1974. *Anarchy, State, and Utopia*. New York: Basic.

O'Connor, Cailin. 2017. "The Cultural Red King Effect." *Journal of Mathematical Sociology* 41 (3): 155–71. https://doi.org/10.1080/0022250X.2017.1335723.

O'Connor, Cailin, and Justin Bruner. 2019. "Dynamics and Diversity in Epistemic Communities." *Erkenntnis* 84 (1): 101–19. https://doi.org/10.1007/s10670-017-9950-y.

Ostrom, Vincent, and Elinor Ostrom. 1965. "A Behavioral Approach to the Study of Intergovernmental Relations." *American Political Science Review* 359:137–46.

Ostrom, Vincent, Charles Tiebout, and Robert Warren. 1961. "The Organization of Government in Metropolitan Areas: A Theoretical Inquiry." *American Political Science Review* 55:831–42.

Pincione, Guido, and Fernando R. Tesón. 2006. *Rational Choice and Democratic Deliberation: A Theory of Discourse Failure*. New York: Cambridge University Press.

Popper, Karl R. 2013. *The Open Society and Its Enemies*. Princeton, NJ: Princeton University Press.

Prinz, Jesse. 2008. *The Emotional Construction of Morals*. New York: Oxford University Press.

Quong, Jonathan. 2011. *Liberalism without Perfection*. New York: Oxford University Press.

Rawls, John. 1974. "The Independence of Moral Theory." *Proceedings and Addresses of the American Philosophical Association* 48:5–22.

Rawls, John. 1980. "Kantian Constructivism in Moral Theory." *Journal of Philosophy* 77 (9): 515–72.

Rawls, John. 1995. "Political Liberalism: Reply to Habermas." *Journal of Philosophy* 92 (3): 132–80.

Rawls, John. 1996. *Political Liberalism*. New York: Columbia University Press.

Rawls, John. 1999a. "Reply to Alexander and Musgrave." In Samuel Freeman, ed., *Collected Papers*. Cambridge, MA: Harvard University Press, 232–53.

Rawls, John. 1999b. *A Theory of Justice*. Rev. Cambridge, MA: Belknap Press.

Rehfeld, Andrew. 2011. "The Concepts of Representation." *American Political Science Review* 105 (3): 631–41. https://doi.org/10.1017/S0003055411000190.

Rehfeld, Andrew. 2017. "On Representing." *Journal of Political Philosophy* 26 (2): 216–39.

Sabl, Andrew. 2012. *Hume's Politics: Coordination and Crisis in the "History of England."* Princeton, NJ: Princeton University Press.

Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.

Scanlon, T. M. 2014. *Being Realistic about Reasons*. Oxford: Oxford University Press.

Schotter, Andrew. 2008. *The Economic Theory of Social Institutions*. Cambridge: Cambridge University Press.

Schumpeter, Joseph. 1994. *Capitalism, Socialism and Democracy*. 5th rev. ed. London: Routledge.

Southwood, Nicholas. 2010. *Contractualism and the Foundations of Morality*. Oxford: Oxford University Press.

Sparrow, Robert. 2016. "If People Were Movies? Free Speech and Free Association." *Journal of Political Philosophy* 24 (2): 227–44.

Sunstein, Cass R. 2002. "The Law of Group Polarization." *Journal of Political Philosophy* 10 (2): 175–95. https://doi.org/10.1111/1467-9760.00148.

Thoma, Johanna. 2015. "The Epistemic Division of Labor Revisited." *Philosophy of Science* 82 (3): 454–72. https://doi.org/10.1086/681768.

Thompson, Abigail. 2014. "Does Diversity Trump Ability? An Example of the Misuse of Mathematics in the Social Sciences." *Notices of the American Mathematical Society* 61 (9): 1024–30.

Thrasher, John. 2014. "Uniqueness and Symmetry in Bargaining Theories of Justice." *Philosophical Studies* 167 (3): 683–99.

Thrasher, John. 2019. "Constructivism, Representation, and Stability: Path-Dependence in Public Reason Theories of Justice." *Synthese* 196 (1): 429–50. https://doi.org/10.1007/s11229-017-1488-7.

Thrasher, John, and Kevin Vallier. 2015. "The Fragility of Consensus." *European Journal of Philosophy* 23 (4): 933–54.

Thrasher, John, and Kevin Vallier. 2018. "Political Stability in the Open Society." *American Journal of Political Science* 62 (2): 398–409.

Tiebout, Charles M. 1956. "A Pure Theory of Local Expenditures." *Journal of Political Economy* 64 (5): 416–24.

Weisberg, Michael. 2007a. "Three Kinds of Idealization." *Journal of Philosophy* 104 (12): 639–59.

Weisberg, Michael. 2007b. "Who Is a Modeler?" *British Journal for the Philosophy of Science* 58:207–33.

Weisberg, Michael, and Ryan Muldoon. 2009. "Epistemic Landscapes and the Division of Cognitive Labor." *Philosophy of Science* 76 (2): 225–52.

Zollman, Kevin J. S. 2009. "The Epistemic Benefit of Transient Diversity." *Erkenntnis* 72 (1): 17–35. https://doi.org/10.1007/s10670-009-9194-6.