

Moral Constitutions

By Gerald Gaus

WHAT APPEARS TO BE A DEEP DISPUTE BETWEEN ANALYSES OF MORALITY may be no dispute at all. Two philosophers may appear to advance radically different theses in moral philosophy (say, different views of justice), which may generate heated debate among commentators — and yet, because each philosopher is asking different questions, and engaging in a different mode of moral inquiry, the dispute may be largely illusory. I suspect that this may be true of many (apparent) disputes between John Rawls and his critics. Throughout his career, but perhaps most noticeably from the mid-1970s onwards, Rawls developed modes of moral and political inquiry distinct from normative moral and political philosophy as it has traditionally been understood. In this essay I am especially concerned with Rawls' short-lived project in what he called "moral theory" — a project that lasted from approximately the mid-1970s to the early to mid-1980s. I believe that the radical nature of Rawls' project, and how fundamentally it departed from moral philosophy as traditionally conceived, has not been properly appreciated. Consequently, many who take themselves to be Rawls' protagonists — for example by advancing certain views of the nature of justice or moral truth — are addressing different questions, employing different modes of inquiry, and so are not offering contradictory theories.¹ And many of those who take themselves to be defending Rawls are actually reinterpreting his work in a way that makes him a much more traditional — and in my view a far less important — philosopher.

My concern, however, is not simply to understand better Rawls' great work. The mode of inquiry characterizing my recent *Order of Public Reason* is basically what Rawls called "moral theory;" I wish to clarify what the book seeks to achieve, and what questions it puts aside. More importantly, however, I wish to suggest that once we clearly grasp the nature of "moral theory" (which I shall be employing as a term of art) as opposed to moral philosophy, we shall better appreciate how the project on moral theory that Rawls sketched can be completed

Gerald (Jerry) Gaus is the James E. Rogers Professor of Philosophy at the University of Arizona and the current director of the university's major in Philosophy, Politics, Economics & Law. His primary areas of study lie in public reason, social and political philosophy, normative ethics, political economy, and philosophy & economics. His most recent book, *The Order of Public Reason* (Cambridge University Press, 2011), draws on aspects of game theory and social evolution to show how respect for the moral freedom of all is possible in the face of social moral indeterminacy. He is currently working on a book, *The Tyranny of the Ideal*, which explores the role of political philosophy in society and the characteristics of societies that encourage increasingly just arrangements.

along the lines that I propose in .

1. Moral Theory As A Mode Of Inquiry

LET ME BEGIN WITH A KEY PAPER OF RAWLS', THE RADICAL MESSAGE OF WHICH HAS NOT been sufficiently appreciated: his American Philosophical Association address on "The Independence of Moral Theory." One message of that paper is clear, and I believe has had reasonably wide uptake — that moral theory is independent of epistemology, the theories of meaning and of language, and the philosophy of mind. What is less often appreciated is what Rawls takes the task of moral theory to be, and how it fundamentally differs from moral philosophy as it is generally conceived of. Rawls writes:

Moral theory is the study of substantive moral conceptions, that is, the study of how basic notions of the right, the good, and moral worth may be arranged to form different moral structures. Moral theory tries to identify the chief similarities and differences between these structures and to characterize the way that they are related to our moral sensibilities and natural attitudes, and to determine the conditions they must satisfy if they are to play their expected role in human life.²

Rawls explicitly stressed that the perspective of the moral theorist must be divorced from that of one committed to a particular moral conception.³ "One thinks of the moral theorist," wrote Rawls, "as an observer, so to speak, who seeks to set out the principles of other people's moral conceptions and attitudes."⁴ As moral theorists, Rawls writes, "we are investigating an aspect of human psychology, the structure of our moral sensibility." Moreover, moral theory is necessarily concerned with social theory, and the feasibility of the sort of society a moral conception instructs us to seek.⁵ In contrast, when taking up the perspective of one devoted to a moral conception one may regard it as the correct theory, as that which provides the definitive account of normativity, justice, or the nature of objective right and wrong.⁶ When advocating a moral conception one may find it puzzling how others can fail to grasp its insights, whereas Rawls repeatedly and approvingly cites Sidgwick's remark in the preface to the first edition of *The Methods of Ethics* that he sought "to put aside temporarily the urgent need which we all feel of finding and adopting the true method of determining what we ought to do; and to consider simply what conclusions can be rationally reached if we start with certain ethical premises." The aim, said Sidgwick, is to adopt "a neutral position" when evaluating the various methods of ethics.⁷

Thus "Kantian Constructivism in Moral Theory," the essay that directly follows "The Independence of Moral Theory," most emphatically does not present a Kantian constructivist account of the true, or philosophically soundest, understanding of normativity or the foundation of the correctness of moral claims.⁸ To take Rawls' constructivism as pursuing the same project as, say, Korsgaard's, is to confuse moral theory as the study of coherent and feasible moral structures with the more traditional project of grounding ethics and normativity on a secure foundation.⁹ Rawls' ultimate aim is to construct a reasonable moral structure — one that coheres with our natural sentiments, our moral psychology, and our understanding of social theory. To object, as does Cohen, that Rawls brings in

criteria exogenous to moral truth or justification, such as publicity, feasibility and coherence with our sentiments, is simply to confuse moral theory with a traditional project in moral philosophy that Rawls sets aside — not as unworthy or misconceived, but as different from his.

“Kantian Constructivism in Moral Theory” has at least two ambitions. Certainly one is to show Sidgwick’s error in dismissing Kantianism as a genuine method of ethics, or in Rawls’ terms, a reasonable moral conception.¹⁰ Rawls shows how the Kantian view can be understood as a reasonable moral structure, constructed out of the conception of persons and social cooperation inherent in the culture of liberal democracy. But Rawls presses further: the nature of a Kantian constructivism, and in particular its emphasis on publicity, is such that it is especially suited to serve as what he calls the “public moral constitution” of our society.¹¹ Although this more daring ambition presses beyond the impartial stance of the moral theorist described in the essay on “The Independence of Moral Theory,” it still abjures the quest to justify a best philosophical account of normativity or the right. Rather, the aim is to employ moral theory’s comparative method to identify a moral structure that best coheres with our moral sensibilities and natural sentiments, and which, given what we know about social theory, can provide the basis of a public and stable moral order. Again, the comparison with Sidgwick is illuminating. *The Methods of Ethics* is not (as some mistakenly read it) a failed philosophical case for utilitarianism: it is an exercise in moral theory. Nevertheless, Sidgwick’s conclusion emerges — overall, he views utilitarianism as the most reasonable moral conception. It is perhaps this feature of both Sidgwick’s and Rawls’ moral theorizing — that out of moral theory arises a view of the most reasonable overall structure — that confuses so many readers, and leads them to mistake their inquiries as a form of normative ethics or metaethics. Thus at least in “Kantian Constructivism” Rawls believed that moral theory can provide a public moral constitution. I return to his proposal in section 3.4.

2. The Moral Constitution

RAWLS’ SETTING ASIDE OF MORAL PHILOSOPHY, UNDERSTOOD AS AN INQUIRY INTO moral truth or the nature and basis of normativity, and focusing instead on moral theory, was a crucial advance in understanding how we might come to uncover or construct a public framework for moral life. Yet, I believe, Rawls’ own proposal was not up to the task. To see why this is so, we need to inquire more deeply first, into the idea of a moral constitution, and then into a constitution that meets the compelling requirement of publicity. A moral constitution that meets the constraint of publicly I will call, following Rawls, “a public moral constitution.”

Now a political constitution is the basis of legitimate and mutually recognized political authority among those who favor competing political views — conservatives, democratic socialists, liberal egalitarians, greens, feminists, libertarians and so on. As philosophers and social theorists have rediscovered in recent years, however, social life is not only ordered through the political, but also by the moral norms (or, we say, social-moral rules), which ground moral authority relations.¹² Social cooperation and, more generally, social life depends on a network of rules that set down the terms for a wide range of social interactions. Having a shared understanding of these terms is the warp and woof of our life

together. These rules do not only coordinate our action—they assure each of us certain claims and spheres of personal authority, as well providing public and definitive terms by which we resolve disputes. No human society has managed to survive without some such framework; the framework characteristic of our trans-national social order is perhaps the most wide-ranging shared system of norms in history.

Taking up the perspective of moral theory, with its reliance on social theory and its concern with feasibility, we can observe that the moralization of this framework is also fundamental to human social existence. Social rules that are understood as moral rules are not simply accorded a sort of overridingness from the social point of view. Because they are internalized by members of the order, they provide internal guidance, which resists temptations to defect and cheat. As moral psychologists have discovered, a critical stage of moral development—one that occurs around eight years of age—is when a person feels guilt over her transgressions, even when they remain undiscovered by others.¹³ And when others do give into these inevitable temptations, those who conceive of the rules as moral requirements and prohibitions take it upon themselves to detect, criticize, blame and sanction the cheaters, greatly stabilizing the social order.¹⁴

These matters are relevant because, as Rawls stresses, moral theory is centrally concerned with the social role of our moral structures and natural sentiments. To be sure, from the perspective of some types of moral inquiry the social role of our moral notions is not relevant to moral justification. What is important to these modes of inquiry is that they help us to see the moral truth, or the nature and grounding of normativity, or vindicate certain views about justice that we reflectively believe to be correct. These questions are set aside by the moral theorist, who turns his attention to our moral practices. His concern is to understand how these moral networks match up with and support our moral emotions and sentiments, and how well they perform the tasks that a society’s moral order must accomplish: the coordination of behavior and the adjudication of claims.

It is appropriate to term this network of moral rules a “moral constitution” as it provides a commonly recognized basis of our authority over each other. When advancing claim on each other, the participants invoke, and recognize, the authority of these rules; the core of our moral constitution is the mutual recognition of this authority. To grasp this interpersonal nature of this authority we must switch from the perspective of the moral theorist, who takes an impartial view of moral structures and the findings of social theory, to the perspective of participants, who view the rules from, as it were, the inside. From this transactional, participant, perspective, our shared moral rules determine when I have an authoritative claim on your action and life, and when you possess such authority over me. As P.F. Strawson observed, we must take up (though never at the same time), both an objective perspective on our social morality and a participant perspective.¹⁵ As participants we see ourselves enmeshed in a system of moral authority over each other that is intimately tied to our natural and moral sentiments. The idea of ourselves as agents is deeply informed by our participation in this public moral framework. Just as our political self-understanding is bound up with our conception of ourselves as citizens, our social identity is bound up

with our conception of ourselves as moral persons in this basic framework. It is not a conception that we are argued into; as young children we naturally take it on. Those few who do not, such as the psychopath or perhaps some types of fanatics, live among us but do not understand themselves as subject to the moral constitution. Whatever our moral stance should be towards these others — and I shall consider later (§5) the moral stance we might take toward them — they are not participants in our system of mutual moral authority.

We now come to the critical problem. The moral theorist, as it were, analyzes this moral constitution from the outside: she sees it from the perspective of moral psychology and social theory, seeking to understand it, and to arrive at standards to evaluate it. As Sidgwick says, she seeks a neutral perspective. But internal to the practice is an acceptance of relations of mutual authority and accountability. The critical question that confronts the moral theorist is the relation between her account of the practice from the outside and the internal participant perspective. Some analyses from the outside, we shall see, must remain hidden from the participants insofar as, if participants accepted such an external analysis, their internal view of the practice would be undermined. Such moral theories would be *self-defeating under conditions of publicity*: the full knowledge that the practice rested on those grounds would undermine the participants' internal view of it. If they accepted that as the basis of the practice, the practice could no longer be maintained by them. In contrast, a moral theory that can be endorsed by the participants — one that is public and transparent to all¹⁶ — is self-sustaining: coming to know its external justification reinforces the internal perspective of the practice.

Now it might be argued that a requirement of publicity is exogenous to the idea of justice.¹⁷ Justice, one might say, is justice, and not publicity or stability. If, however, our concern is a moral constitution that sustains the relations of normative authority on which our social order rests, publicity is necessary given what Rawls calls "the social role of morality."¹⁸ A *public* moral constitution is one for which participants, understanding its objective (external) analysis and justification, may rationally affirm its normative status in their interpersonal dealings. The normative force of such a moral constitution is consistent with the free use of the reason of each. As participants come to reflect on their shared moral life, they each come to affirm its normative force in their dealings with each other, and so their shared social life is put on a firm footing. The more they reason, the more the normative basis of their life together is affirmed. In one sense of the term, such a society is well-ordered. If, in contrast, on the view of each the normative force of the basic moral framework depends on not looking too deeply into it, or requires entertaining distorted ideas about its basis and purposes, then like the constitutions of the *ancien régime*, its legitimacy will crumble when exposed to the light of actual human reasoning. As David Gauthier observed, if in the end our moral constitution depends on claims that we find beyond credibility — to endorse its normative authority one must also endorse supernatural or non-natural entities or rarified philosophical constructions — it will face a foundational crisis.¹⁹ Like religion, it will not stand up to critical public scrutiny. Although the theologians and moral philosophers may assure us that we would really see how it is justified if only we could appreciate the form of justice or intuit the relevant non-natural

qualities, the mass of real human reasoners will be, at best, left unmoved. Or worse, they may simply be amused by supposed insights of the philosopher.

3. Four Inadequate Accounts Of A Public Moral Constitution

GIVEN THIS CRUCIAL PROBLEM LET US CONSIDER FOUR ANALYSES OF OUR MORAL constitution, and why they fail to be self-sustaining given publicity: the merely positive, the instrumentalist, that of moral philosophy, and Rawls' Kantian constructivism in moral theory.

3.1 The Positivist Analysis

The merely positive stance accepts, as do all tenable views, the necessity of a moral framework for social life, but insists that moral theory can be *only* an empirical, sociological, study of social norms, having no normative force — no real normative standing. What people take as normative and what is truly normative, we are told, are simply different questions, and it does no good to blur them as the moral theorist seeks to do. The positivist view commits to an external analysis that appreciates the social benefits of the framework, but abjures any normative commitments. It is the view of the pure social scientist. Moral psychology and social theory are one thing, normative inquiry quite another.

Empirical analysis certainly has a great deal to teach; however, it cannot serve as the basic public understanding for the framework. As an explanatory perspective it is immensely edifying, but if we suppose that is *all* there is to say from the objective, external view, the positivist analyses is transformed into a debunking view. In this regard it is reminiscent of the classical theories of ideology, which sought to replace philosophical argument with a purely scientific study of ideas.²⁰ In the end, the claim of Karl Mannheim and Karl Marx was that they saw *through* the pretensions of normative political discourse to see things as they really are — for Marx, to see the discourse as a form of struggle between class interests.²¹ Participants cannot endorse this view of their practice and yet carry on with it as it was, for the objective analyses shows the internal perspective to express a false consciousness. Participants in the moral order understand themselves to be making moral demands, to be claiming moral authority over others, and experience the moral emotions in the face of moral failure. If that all turns out to be an illusion — they come to see that there is no normative basis for their practice — the participants cannot sustain their practice, except as a sort of wishful thinking, as some older children profess to believe in Santa Claus in order to make Christmas more enjoyable.

3.2 The Instrumentalist Analysis

Gauthier believed that that morality's foundational crisis can be resolved only by grounding our moral constitution in prosaic, instrumental, justification. After all, every plausible analysis acknowledges that a well functioning moral constitution is necessary to social life: perhaps this necessity is all the normative ground we need. All reflective rational people will see its benefits, and so see its normative necessity. Yet this grounding of morality from the objective point of view remains unable to serve as a public basis for our practice of moral authority, for it too is unable to sustain the internal, participant, perspective under conditions

of publicity. That it is useful to observe a practice in which the participants acknowledge each other's authority to advance moral demands does not establish that they possess such normative authority, any more than showing that it would be useful to treat the laws of one's government as authoritative makes them authoritative. Even if there were no such thing as legitimate authority relations it still might be useful indeed for us all to treat the state's edicts *as if* they had authority. Or, to take another case: that it is socially beneficial to have a practice of aesthetic evaluation cannot be the grounding reason why Frank Lloyd Wright's *Falling Water* is a superlative work, nor can the reason be that it is socially useful to see it as superlative.

As Strawson indicates, the instrumentalist gives the wrong sort of normative basis of our moral demands on each other.²² The instrumentalist takes the objective perspective, and tells us that our practice is valuable, but that value cannot be what here and now grounds my moral demand that you accord me my due. When I make a moral demand on you to act, I do not claim that you should perform the act simply because it is good for you (or for society as a whole) that you so act, but because you owe it to me to so act.²³ I am apt to resent your failure to act, and I will blame you for ignoring my demand that you do as you ought. To be told that the public basis of my claim is simply that it is useful for participants to think, feel and act in these ways cannot be endorsed from within the practice as its sole grounding.

In his early essay on "Two Concepts of Rules" Rawls denied this. Utilitarianism, he argued, could serve as a justification for our practice of punishment. The utilitarian principle could, from the outside, justify the practice of punishment, but from within this practice we would take the fact that a person violated certain rules as necessary and (with some added details) sufficient for other participants in the practice to have the authority to punish him, and presumably to possess the reactive attitudes of resentment and indignation, and for the wrongdoer to feel guilt and remorse. The utilitarian principle could be the sole justification of the practice, and within the practice there would be "moral offices" and authoritative relations that are not empowered to directly consult the utilitarian standard.²⁴ While Rawls was correct that a utilitarian could design such a practice and approve of its operation, the participants could not acknowledge utilitarianism as the sole public normative basis of the claims they make within the practice.²⁵ The normative fact that the sole basis of all they do and feel is that it is socially useful for them to do and feel it must remain insulated or hidden from their activity within the practice. If I know that the normative grounds of my punishing you is not that you have cheated me, or harmed an innocent, but that by so doing you have violated the rule of a practice that has the sole aim of promoting the social welfare, I no longer see the basis of the punishment in a transactional wrong — a manifestation of lack of regard to me or another — and if that is so, then my moral sentiments must themselves be transformed. "Bad person," I might say, "you have violated a socially useful norm, and it is socially useful for me to be angry at you!" As Strawson reminds us, "What is wrong is to forget that these practices, and their reception, the reactions to them, really are expressions of our moral attitudes and not merely devices we calculatingly

employ for regulative purposes. Our practices do not merely exploit our natures, they express them."²⁶ Our practices express our sentiments that we are owed a certain regard by those with whom we interact. The instrumentalist and utilitarian can explain the usefulness of our moral sentiments and attitudes in regulating social life, but the public realization of their usefulness is not enough to ground our reactive attitudes.

3.3 Moral Philosophy and the Normative Perspective

Moral philosophy understood as an account of the moral truth, the grounding of normativity, or a preferred justification of moral prescriptions, appears to provide what the positivist and instrumentalist analyses lack, but which a public moral constitution must possess: a public account of the normativity of the practice that makes sense of its internal commitments and presuppositions, and its transactional claims and responses. If our practices track moral right and wrong, duty and obligation, then we appear to have finally resolved our problem. We can understand from the outside the usefulness of the practice and the constituent normative structures, while accounting for the internal perspective, since its usefulness at least is consistent with (and perhaps supervenes on) its correct transactional morality. As participants our moral sentiments and judgments are engaged when others do what is wrong or harmful. If the justification of our practice tracks, or at last coheres with, moral truth, or its rules and requirements are genuinely grounded in the normative, participants can embrace this as the self-sustaining justification of their moral constitution. We thus seem to have finally arrived at a truly public moral constitution.

Alas, as Rawls showed so clearly, given the facts of reasonable pluralism, no moral philosophy thus understood can serve as the public grounding of our moral constitution. We must not interpret Rawls' doctrine of reasonable pluralism as applying only to comprehensive conceptions of the good; he is entirely explicit that the "burdens of judgment" apply to notions of the right.²⁷ Because moral issues are so complex, because there are multiple values that must be weighed and there is no clear way to weigh them, because the empirical evidence is uncertain and what we take as evidence is often dependent on our previous beliefs and values, we disagree about the correct moral philosophy, be it divine-command grounded, Kantian, natural-rights based, intuitionist, virtue-based, moral sentimentalist, conventionalist, or whatever. The public knowledge that, say, a certain form of Kantianism (constructivist or otherwise) is *the* normative grounding of our moral constitution could not be endorsed by the mass of good-willed and rational participants. If the practice hinges on *that* being the case, then in the eyes of many the practice will, paradoxically, lose its normative force in their eyes. For whether or not the Kantian is the philosophically best case, many good-willed and reflective people believe it is not the best or correct view, and so the public acceptance of it as foundational to our moral practice will lead to the constitution's loss of legitimacy for them. They cannot go on with their participation in light of the fact that, as they see it, the entire edifice is grounded on an error. On their understanding, their claims are now without normative force. And so they cannot accept the moral practice as genuinely authoritative, nor can they endorse the authority of the demands advanced by others on the

basis of the rules of the practice.

This is why, as Rawls taught us, appeal to the moral truth, or the best philosophical account of morality, is of no help in sustaining the normative basis of our shared moral life — even if I really possess the truth. In a society that was characterized by consensus on moral philosophy, perhaps grounded on a deeper religious consensus, a moral philosophy may be the basis of a public moral constitution. Or even if, given disagreement on these matters there was a consensus that some group who agreed among themselves had superior insight and so should be obeyed, a moral philosophy could serve as the public basis of their moral constitution. But we do not live in this world. Given reasonable pluralism, there is great disagreement about what is best or correct in moral philosophy. Most do not endorse the favored moral philosophy.

To be sure, many moral philosophers believe that they have a moral obligation to always act on the moral truth as they see it.²⁸ As such philosophers see it, their moral demands have authority over others simply because they are normatively correct. Now I do not wish to deny that one of these philosophers may be correct about the best or true moral philosophy or that what such a philosophy might be is an interesting question worth pursuing. But the question for us, here and now, is whether this sort of controversial moral philosophy can serve as the public self-sustaining grounds for a practice of moral authority. I do not believe it can.

In his Lectures on the *Principles of Political Obligation* T. H. Green resists the idea that *de jure* sovereignty is simply “rightful authority” in the sense of an authority based purely on the correct normative basis. Grounding the authority of my claim over others simply of the supposition that it is the normatively correct claim, even though it has no practical force over those on whom I assert it, is like a claim to political authority based on “the mere name of a fallen dynasty exercising no control over men in their dealings with each other.”²⁹ Instead, Green argues, the distinction between *de facto* and *de jure* sovereignty “has natural meaning in the mouths of those who, in resisting some coercive power that claims their obedience, can point to another determinate authority to which they not only consider obedience due, but to which obedience in some measure is actually rendered.”³⁰ Green’s general point, and he seems entirely right, is that a political authority that has, or is likely to have, no practical effect is no political authority at all, as it cannot perform its main task of sorting out actual disagreements and coordinating activities. To be any sort of authority at all there must be some general recognition of it; only then can it perform its designated tasks. If it is not generally recognized as an authority, we might argue that it *ought* to be an authority, that it alone *could qualify* as a justified authority, but we cannot claim that it now is such an authority. A crucial job of moral authority is to regulate and coordinate social interaction by invoking the moral sentiments; if so, an authority that is not recognized simply is unable to perform the office of an authority, as one who is not socially recognized as a leader is unable to fulfill the position of “group leader.” We can say that a person who is not recognized — either explicitly or implicitly — as a leader *ought* to be the leader, that she would be a wonderful leader, or even that she would be the only leader worth having; but we cannot say that a person no one pays attention to is *now* the leader. And neither can we say that a

philosophical account of normativity to which few pay attention grounds the *de jure* system of moral authority that is a justified moral constitution.

3.4 Kantian Constructivism in Moral Theory

We can now see the deep attractions of Rawls’ Kantian constructivism as a moral theory. If we wish to uncover a public basis of our moral constitution, we should begin with the materials that are already part of that constitution: the models of the person, the structure of the natural and moral sentiments, and so on. By building up our normative account from, as it were, the inside of the practice, we seek to ensure that the ideas we employ are already commitments of the participants rather than external standards of correctness or truth. Thus, as Rawls tells us,

The “real task” of justifying a conception of justice is not primarily an epistemological problem. The search for reasonable grounds for reaching agreement rooted in our conception of ourselves and our relation to society replaces the search for moral truth interpreted as fixed by a prior and independent order of objects and relations, whether natural or divine, an order apart and distinct from how we conceive of ourselves. The task is to articulate a public conception of justice that all can live with.... What justifies a conception of justice is not its being true to an antecedent order to and given us, but its congruence with our deepest understanding of ourselves and our aspirations, and our realization that, given our history and the traditions embedded in our public life, it is the most reasonable doctrine for us. We can find no better basic charter for our social world.³¹

When Rawls tells us that moral theory replaces the idea of the moral truth with that of the “reasonable,” he is not simply informing us that he prefers a term for “the best moral philosophy” that does not have strong realist connotations. The idea is much more radical: moral theory simply is not concerned with *any* understanding of the philosophically best normative account, rather, it seeks a “basic charter for our social world” that we can all appreciate as a public basis of our normative social relations.

It is difficult to exaggerate what a radical and insightful project Rawls launched — one that connects moral theory to many of the traditional problems of political philosophy. Yet, as we all know, the phase of his project characterized by the two key essays on which I have focused, “The Independence of Moral Theory” and “Kantian Constructivism in Moral Theory,” was relatively short-lived — roughly the mid-1970s to the early-to-mid-1980s. By the mid-1980s the project was turning into political liberalism.³² I do not wish to engage in intellectual biography, so I am not concerned with why Rawls gave up on it as a general approach to our moral constitution, but I do think we can see a deep problem.

As we saw regarding both Sidgwick and Rawls (§1), from the neutral and impartial perspective of moral theory, each hoped that one moral conception or method would emerge as clearly the most adequate. But the very project of putting “aside temporarily the urgent need which we all feel of finding and adopting the true method of determining what we ought to do” and considering “simply what conclusions can be rationally reached if we start with certain ethical premises” is in tension with advocating one method as the impartially best — its premises and

structure most reasonable. Instead of, as Rawls put, examining “other people’s” reasonable moral conceptions (§1), the moral theorist comes to advocate his favored conception as the most reasonable. The insuperable problem, though, is that given reasonable pluralism, no single moral conception, understood as a more or less complete structure of one’s moral commitments that coheres with one’s sensibilities and natural sentiments that is socially feasible and public, will be seen as most reasonable by good-willed and reflective moral agents. For a moral conception to serve as the public normative basis of our moral constitution, at least if it is to be the sort of stable and public moral constitution Rawls sought in “Kantian Constructivism,” there must be wide acceptance among the public = that it provides “sufficient reason” to take the moral practice and its constituent transactional relations as justified.³³ It is imperative to distinguish (a) holding that some subset of normative considerations provide a *pro tanto* justification of a practice from concluding that (b), given one’s moral and evaluative perspective, the practice is, all things considered, normatively justified. If we only have identified some reasons, participants as actual persons with their normative views may well conclude the practice does not have sufficient justification, and hence deny its authority. If our aim is a stable moral constitution to which participants “grant” moral authority, it thus will be no avail to bracket some of their normative considerations, and insist that they do not count; one cannot simply insist that only the ideas identified by the Kantian conception, or only so-called liberal values, are relevant. Participants in the practice will rightfully consult their complete understanding of the most compelling standards of evaluation when determining whether the claims of the practice have the requisite moral force.

Thus to justify the moral authority characterizing our moral constitution, we must not simply point to *some* legitimizing reasons, but to ones that are sufficient to legitimate the claim to authority. And something like this must be the case if the moral constitution is to be self-sustaining: when you claim authority over me and demand that, for example, I cease pushing you on the subway, it must generally be the case that, my own reasoning confirms this, and so I comply. As Mill stressed in *Liberty*, and as recent empirical work has confirmed, social morality certainly invokes sanctions for noncompliance;³⁴ yet it also is the case that, compared to the legal system armed with its extensive executive powers, mass voluntarily compliance with the moral constitution is necessary for a stable social world. As Rawls appreciated, such stability requires a generalized Nash equilibrium: conformity to the moral constitution must generally be the best response to the conformity of others.³⁵ But this means it is not enough that we all endorse some of the elements of Kantian constructivism as a moral theory (such as its conception of a moral person); we must, as it were, hold that the construction best explicates our overall moral outlook, at least insofar as we are concerned with social morality.³⁶ In short, we must all embrace Kantian constructivism as the preferred public moral conception grounding our moral constitutional consensus.

To be sure, not all moral practices can be given a public normative foundation, for some are oppressive and at least some participants reasonably deem them without normative force. The question is not whether an adequate account of the public normative basis of moral practices will sometimes deem a practice as lacking normative force, but whether any moral practice could be

deemed to have moral force if such force requires consensus on a single moral conception. If that type of consensus is required, in our world of reasonable pluralism we shall be unable to equilibrate on a basic charter for our social world.

4. Convergent Normativity:

Moral Constitution Making From The Ground Up

RATHER THAN COMMENCING BY IDENTIFYING THE MOST REASONABLE BASIS FOR OUR moral constitution, we might instead think of constitution making as process that is built from the ground up.³⁷ Each good-willed, reflective and rational individual, employing the evaluative standards that she views as most compelling, seeks to determine whether she has her own normative reasons to endorse and internalize the rules of our practice. We might, then, view the moral constitution as arising from interacting individuals, each consulting her own normative standards and taking into account the actions of others, acting on moral rules that she deems to have moral force. Rather than the American, we would do better to think of the British Constitution and the Common Law; instead of a definitive constitutional choice situation, we should instead think of an evolutionary process in which participants come to converge on a public moral constitution. This idea is common to moral thinkers as diverse as classic philosophers such as Hume and Hegel, and more contemporary thinkers such as Peter Strawson, Kurt Baier and F.A. Hayek.³⁸ On this view, as Baier put it, “a society’s morality is the joint product of the moralities of its individual members. As far as its content is concerned, individual members are its joint makers, not merely its subjects.”³⁹ In Rawls’ terms, the constitution will be publicly justified because each sees it as the object of an “overlapping consensus.”⁴⁰

An adequate theory of the moral constitution, then, will inquire, for any actual or proposed rule of justice, whether good-willed and rational agents, reflecting on their entire set of normative standards, have sufficient reason to internalize the rules as authoritative and adopt the moral attitudes. If they do — if given one’s set of evaluative considerations, which sometimes overlap with those of others and sometimes are distinctive — then in one’s view the rule has passed the test of normativity: it is a rule that one has reason to treat as is befitting a rule of the practice of morality. There is no reason to think that a person only has reason to adopt these sentiments and attitudes towards the rule she thinks best, as dictated by her moral philosophy or religion. As the politically mature citizen does not require that only the political constitution she deems ideal can ground legitimate authority over her, the morally mature participant in our moral constitution does not insist that only her preferred moral conception can ground interpersonal moral claims. In reflecting on the normativity of our moral constitution we are not concerned with bargaining, bluffing or disputes in the pages of a philosophy journal. The question on which one must focus is whether, given her own understanding of normativity and value, does this understanding endorse these rules as having normative authority? Given the pros and the cons from one’s own moral conception, would some basic social charter for our social world — a certain practice of morality — be worthy of endorsement? Simplifying — but not so much as to obscure the critical point — if it is, a person has reason to embrace the moral constitution (or constituent rule) as within the range of

acceptable charters for our social world.

Of course different people will apply this test with different results. Some will identify a wider range of acceptable rules (or charters for our social world), while for others it will be narrower. Suppose each good-willed person, seeking to live under a shared moral constitution with others, evaluates the options; let us call *S* the eligible set of moral constitutions that all deem to be worthy of endorsement. We are faced with three possibilities: *S* will be a (i) singleton, (ii) it will be null, or (iii) it will have multiple elements.

(i) The aim of much of the social contract tradition has been to show that the eligible set contains just one element or that we all agree what is the best option. Rawls' *Theory of Justice* explicitly sought to demonstrate that all reasonable citizens would rank as best his two principles of justice: in a pairwise comparison with a list of alternatives each would agree that Rawls' principles are best.⁴¹ There would thus be no reason to settle for any alternative: our reason leads us to agree. As is well known, Rawls' later writing evinces increasing skepticism that we will converge on a single best conception: as I have pointed out, the burdens of judgment apply to the right as well as the good.

(ii) But if the use of our reason leads us to disagree about the right, the second possibility looms — we face the quintessentially modern worry that perhaps no moral constitution is acceptable to all. Many cannot assuage their anxiety that in our contemporary world of normative diversity a public moral constitution is impossible, and so public moral life must be simply a sustained moral conflict, or the rule of the wise over their benighted fellows. Can a diverse population, torn by basic conflicts about religion and value, nevertheless freely endorse a common moral constitution, or does an effective basic charter for our social world necessarily rely on extensive use of force and duplicity, by which the powerful impose their favored scheme on the rest? The liberal tradition's battle against the pessimistic conclusion has been grounded on two fundamental insights. As Hobbes taught us, a shared framework for social order is a great good to all; a wide range of values, ends, and lives based on different convictions about our place in the universe are furthered by a cooperative, peaceful, social order. To be sure, those whose life's interest is in dominating others will not find their interests served by a cooperative social order, and in the end we must have something to say about these individuals. But we must not lose sight of the great good that almost all perceive in a cooperative and stable moral framework for our social existence. In addition to what we might deem the Hobbesian insight, the Millian insight has been that in a social world of great diversity the only possible moral constitutions that we can all endorse are ones that ensure to everyone various spheres in which one's deepest normative commitments hold sway — families, churches, voluntary associations and, in general, a private sphere.⁴² A constitution building on the Millian insight will, of course, be one that grants pride of place to individual liberties, especially those that Benjamin Constant called the "liberties of the moderns:"

the "right of everyone to express their opinion, choose a profession and practice it, to dispose of property, and even to abuse it; to come and go without permission, and without having to account for their motives or undertakings. It

is everyone's right to associate with other individuals, whether to discuss their interests, or to profess the religion which they and their associates prefer, or even simply to occupy their days or hours in a way which is most compatible with their inclinations or whims.⁴³

(iii) If the liberal project is successful in fending off the pessimistic conclusion, we seem left with our third possibility: reasonable people will see the importance of securing their basic interests in bodily integrity. As self-directing agents in the world we certainly have a fundamental interest in securing the basic conditions for our continued agency.⁴⁴ And we have a fundamental interest in some demarcation of what is mine and thine. Of course, about all these issues we disagree on the best or optimal specification of the right; there is always the need to solve the impure coordination problem of selecting one among many possible equilibria to regulate our common life. Understanding moral constitution-making from the ground up requires that we first identify a set of possible norms that all can endorse given their evaluative frameworks (their conceptions of normativity), and then show how we can come to settle on one. As I have argued, finding such a constitution can be understood as a large scale iterated impure coordination problem.⁴⁵ There are a number of moral constitutions that each has sufficient reason to endorse, and we disagree about which is best, but there are great goods that are achieved only when we come to share a common, justified, moral constitution regulating our social existences. Participants seeking a moral constitution can, under a variety of circumstances, converge on a set of rules that, each for her own reasons, will endorse and see as having normative force.

5. The Ins And The Outs Of Public Justification

MANY MISCONCEIVE THE MORAL CONSTITUTION OF LIBERALISM AS ONE THAT EXPRESSES the commitments of a specifically liberal value system, and so only "committed liberals" can endorse a liberal moral constitution. This is, I believe, a fundamental error. A liberal normative framework serves as a public constitution in a diverse society because its rules are those that have evolved to allow us to live with, and in innumerable ways benefit from interchanges with, diverse others. A liberal moral constitution can be endorsed by the normative commitments of Catholics, evangelicals, atheists, conservatives, secular philosophers, Randians, and New Agers. As the economist Paul Dragos Aligica shows, some types of normative institutional arrangements arise *out* of heterogeneity.⁴⁶ If we take heterogeneity of preferences, values and moral commitments seriously right from the outset, we will better appreciate how some moral constitutions arise out of normative diversity whereas others require a "normalization" or homogeneity of basic interests and values, and so will always been unstable in our contemporary world.⁴⁷

One thing can be conclusively demonstrated: for any population *N*, a moral constitution that allows all to draw on their diverse normative frameworks in endorsing it will never do worse, and can do better, in publicly justifying a moral constitution than a view that insists on a preferred moral conception.⁴⁸ That is, moral constitutions based on normative diversity weakly dominate those that depend on a preferred moral conception. The proof is trivial. Either the members

of population N all share the preferred moral conception M or they do not. If they do, then the preferred conception view and the normative diversity account are equivalent. Suppose, then, that N is divided into at least two groups, those that hold the preferred moral conception M and those that hold some competing conception M^* , such they do not endorse M . On the preferred conception view only those who hold M can endorse the public basis of the moral constitution, so only those will see it as authoritative. On the diverse bases account, M and M^* may endorse a common constitution C , which departs from both M and M^* . If they do, we have achieved a wider public basis; if M^* cannot, as it were, be brought on board, then the result is again equivalent to the preferred conception view, which relies only on M . The only reason to insist on the preferred conception view — and this has been explicitly expressed by some⁴⁹ — is that if we restrict the public basis of justified moral constitutions to M , we are thus assured that our constitution must be based on M , whereas if we allow both M and M^* , the ultimate constitution C may be different from either, yet one which draws on the views of proponents of M^* as well as M . If the aim of the exercise is to vindicate only constitutions allowed by M , then the preferred conception view makes sense; if the aim is to show that we can share a public moral constitution under contemporary conditions, then we must adopt the diverse normative basis account. Because it weakly dominates the preferred conception account, it enhances the resources for achieving a public moral constitution.

Of course there will always be limits to the types of heterogeneity that can support a shared moral constitution. The all-too-typical worry starts with some extreme instance — say, the Nazi, the jihadist or the serial killer — and so concludes that, because such persons cannot endorse the moral constitution, there really is no publicly justified constitution that provides sufficient reasons to all. In moral theory it is a grave error to commence with the most puzzling and radical cases. Perhaps they are a reasonable first consideration when evaluating some moral philosophies that advance claims to hold for all possible cases, but moral theory is concerned with the moral structures that we — you and I — typically confront, and the problems that most worry us. Our main problem is whether in a social world characterized by deep diversity of religion, philosophy, culture, and ideology, there can be a moral constitution that, despite disagreements, good-willed, normal moral agents can all endorse as possessing normative force. The question is not whether one can describe a person whose values are so hostile to a cooperative and moral social life that she cannot be part of a diversity-based moral constitution, but just how serious the problem is.

At some point, of course, any moral constitution must draw a line, determining that some person is not good willed (she just wants to kill infidels, or torture, or whatever), or simply has no reason to endorse our social morality (perhaps she is a psychopath). Suppose that we have been able to “draw the line” in such a way that we can say we have two groups, Full Members (of our moral constitution) and Complete Aliens — those who are not in any way members. Between the two, I am supposing, there is no convergence on any social-moral rule (a radical supposition indeed given way a moral constitution helps us solve the enduring problems of social life, but let us proceed with the philosophical

stipulation). With Full Members we can establish a rich set of free moral relations and mutually beneficial interactions, we can sustain the moral sentiments, and we can achieve trust through a set of social-moral rules that all normal participants have reason to endorse and act upon, conditional upon enough others doing so. When I demand that a Full Member must do, or must not do, something, I am asserting a sort of moral authority over her that she has sufficient accessible normative reasons to endorse. And we thus have the foundations for a common practice of moral responsibility.⁵⁰

Now consider the Complete Alien. We must assume that the Complete Alien simply has no sufficient reason to endorse any of our moral practices, and we have insufficient reason to revise ours to accommodate him. With such a person — at least the quintessential case — none of the normal moral emotions, reactions, or presuppositions hold, and trustful, mutually beneficial relations are apt to be impossible. Our relations with the Complete Alien are thrown back onto our own first-person convictions about what is the moral thing to do, convictions that have no authority over him. We must, as it were, act outside the constitution: we must act on our own moral convictions without claim to moral authority, as the Complete Alien will act on his. The moral philosopher may have much to say about this case; perhaps he does not even see it as really distinct from our interactions with those who share our practice. In any event, because it stands beyond the limits of a public charter for our shared social world, moral theory has little to contribute.

6. Conclusion

MY MAIN AIM HAS NOT BEEN TO DEFEND THE LIBERAL NORMS AS THE *BONA FIDE* SOCIAL morality endorsed by the normative convergence of diverse views, but to stress how a certain view of moral theory responds to the failures of other modes of moral inquiry to adequately analyze how we may achieve a public moral constitution — a basic charter of our social world. Moral philosophy is set aside; it has different tasks, but is inadequate to this one. Rawls’ great insight was to see how moral theory may provide the basis for an analysis of our moral constitution. The Kantian constructivist project sought an agreement on a reasonable moral theory as the solution to a common life characterized by deep and enduring metaphysical, ethical and philosophical disputes. There is no reason, however, to suppose that we share the same view of a reasonable normative basis for our shared social world. As he began to see in his crucial move to the idea of an overlapping consensus, a public basis of normativity could be grounded on a public recognition of the convergence of diverse views of normativity on a shared constitution. But this is moral theory, where we set aside concerns with the true or the best account of morality and normativity, in favor of what can be publicly and reasonably endorsed as normative. In moral theory we must build the public moral constitution up from diverse individual views of normativity. ◻

*Philosophy
University of Arizona*

Notes

*Versions of this paper were presented to a workshop on public justification at the Darmstadt Technical University organized by the Normative Orders Research Project, to the University of Rijeka Scientific Colloquium, and to seminars at the Murphy Institute of Political Economy, Tulane University and the PPE group at the University of Pennsylvania, and as the 2012 Brian Barry Lecture and the London School of Economics. My thanks to all the participants, as well as to the members of my graduate seminar on Rawls at the University of Arizona. Thanks too to Dave Estlund, for a great conversation about these matters over a couple of IPAs.

¹ I believe that this is the case with many of G.A. Cohen's criticisms of Rawls' political theory in *Rescuing Justice and Equality* (Cambridge, MA: Harvard University Press, 2008), Part II. Cohen criticizes Rawls for failing to advance a true theory of justice, but only what he calls "optimal rules for social regulation." Cohen, I believe, fundamentally misunderstands the relation in Rawls' thinking between justice and rules for social regulation: the nature of this relation is the chief concern of the present essay.

² John Rawls, "The Independence of Moral Theory" in *John Rawls: Collected Papers*, ed. S. Freeman (Cambridge, MA: Harvard University Press, 1999): p. 286.

³ *Ibid.*, p. 288.

⁴ *Ibid.* Emphasis added.

⁵ *Ibid.*, p. 296.

⁶ *Ibid.*, p. 288.

⁷ Henry Sidgwick, *The Methods of Ethics* 7th ed. Chicago: University of Chicago Press, 1907 [1962], p. vi. See Rawls, "The Independence of Moral Theory," p. 290; "Kantian Constructivism in Moral Theory," in *John Rawls: Collected Papers*, p. 341.

⁸ Cf. the interpretations of Cohen, *Rescuing Justice and Equality*, chap. 7; Ranier Forst, *The Right to Justification: Elements of a Constructivist Theory of Justice* (New York: Columbia University Press, 2012), chap. 2.

⁹ Christine M. Korsgaard, *The Sources of Normativity* (Cambridge: Cambridge University Press, 1996). See also Forst, *The Right to Justification*.

¹⁰ Rawls, "Kantian Constructivism in Moral Theory," p. 342.

¹¹ *Ibid.*, p. 326.

¹² Outstanding contemporary works here are Cristina Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms* (Cambridge: Cambridge University Press, 2006); Ken Binmore, *Natural Justice* (Oxford: Oxford University Press, 2005); Shaun Nichols, *Sentimental Rules* (Oxford: Oxford University Press, 2004).

¹³ I discuss this research in *The Order of Public Reason* (Cambridge: Cambridge University Press, 2011), §11.4.

¹⁴ See *ibid.*, §7.

¹⁵ P. F. Strawson, "Freedom and Resentment," *Proceedings of the British Academy* 48 (1962): 187-211, at p. 189-90.

¹⁶ Cf. Rawls, "Kantian Constructivism in Moral Theory," pp. 322ff.

¹⁷ Cohen, *Rescuing Justice and Equality*, chap. 8.

¹⁸ See, for example, Rawls, "Kantian Constructivism in Moral Theory," p. 342.

¹⁹ David Gauthier, "Why Contractarianism?" in *Contractarianism and Rational Choice*, ed. P. Vallentyne (Cambridge: Cambridge University Press, 1991), pp. 15-30.

²⁰ See further my "Ideology, Political Philosophy, and the Interpretive Enterprise: A View from the Other Side," in *Liberalism as Ideology: Essays for Michael Freedman*, edited by Ben Jackson and Marc Stears (Oxford: Oxford University Press, 2012), pp. 178-98.

²¹ See, for example, Karl Mannheim, *Ideology and Utopia* (London: Routledge and Kegan Paul, 1936); Karl Marx and Friedrich Engels, *The German Ideology, including Theses on Feuerbach* (Amherst, NY: Prometheus Books, 1998).

²² Strawson, "Freedom and Resentment," pp. 189-90; see also Stephen Darwall, *The Second-Person Standpoint* (Cambridge, MA: Harvard University Press, 2006), pp. 15-18.

²³ Note that I am putting aside here the otherwise important distinction between an

instrumentalist justification of our moral practice that seeks to show that it is beneficial to each participant (such as Gauthier's) and one that aims to demonstrate that it collectively beneficial (as does the utilitarian's). In this context they share a similar problem of not adequately grounding the accountability relations that are part of moral obligations. Kurt Baier, I believe, makes a similar point in "Moral Obligation," *American Philosophical Quarterly* 3 (July 1966): pp. 210-226, at 218-222. I am grateful to Stephen Darwall for calling my attention to Baier's analysis.

²⁴ Rawls, "Two Concepts of Rules" in *John Rawls: Collected Papers*, pp. 38-9.

²⁵ It is important to stress that an objective, instrumental, grounding could always be (and indeed must be) a public ground of the practice. As Strawson noted "It is far from wrong to emphasize the efficacy of all those practices which express or manifest our moral attitudes, in regulating behaviour in ways considered desirable; or to add that when certain of our beliefs about the efficacy of some of these practices turn out to be false, then we may have good reason for dropping or modifying those practices." "Freedom and Resentment."

²⁶ *Ibid.*

²⁷ Rawls, *Political Liberalism* (New York: Columbia University Press, 1996), p. xlix.

²⁸ See, for example, the remarks of Richard Arneson and Eric Mack on my account at <http://www.cato-unbound.org/archives/october-2011-for-a-new-liberalism/>

²⁹ T. H. Green, *Lectures on the Principles of Political Obligation*, in *Lectures on the Principles of Political Obligation and Other Writings*, edited by Paul Harris and John Morrow (Cambridge: Cambridge University Press, 1986), §105.

³⁰ *Ibid.*

³¹ Rawls, "Kantian Constructivism in Moral Theory," pp. 306-7. Emphasis added.

³² I consider Rawls's transition from *Theory of Justice* to *Political Liberalism* in more depth in my "The Turn to a Political Liberalism" in *A Companion to Rawls*, ed. J. Mandel and D. Reidy (New York: Wiley-Blackwell, forthcoming).

³³ In his late political writings Rawls indicates that an exercise of authority is only legitimate when one can give "sufficient" reasons for it that others are reasonable to accept. Rawls, "The Idea of Public Reason Revisited" in *John Rawls: Collected Papers*, p. 578. Rawls is talking here about the exercise of political power rather than authority; I assume that the point holds for authority. Indeed, it seems more powerful when applied to authority, for *in extremis* we may have to employ political power without such justification, but when so doing we act without authority. See §5 of the present essay.

³⁴ See Mill, *On Liberty*, in *The Collected Works of John Stuart Mill*, ed. J.M. Robson (Toronto: University of Toronto Press, 1963), vol. 18, chap. 1, ¶5ff. For some evidence, see *The Order of Public Reason*, §§7-8.

³⁵ This theme in Rawls' work has been emphasized by Paul Weithman in *Why Political Liberalism? On John Rawls's Political Turn* (New York: Oxford University Press, 2010); compare *The Order of Public Reason*, §17.

³⁶ On the idea of an explication of our moral convictions, see Rawls, "Outline of a Decision Procedure for Ethics" in *John Rawls: Collected Papers*, pp. 77ff. This idea is at the heart of Rawls' method reflective equilibrium which is to organize one's moral convictions in way that one can affirm "with conviction and confidence." "The Independence of Moral Theory," p. 289.

³⁷ I owe this way of putting the point to Pete Beotke, "Constitution Making From the Ground Up." See <http://www.cato-unbound.org/2011/10/20/peter-j-boettke/constitution-making-from-the-ground-up/>

³⁸ See Kurt Baier, *The Rational and the Moral Order: The Social Roots of Reason and Morality* (La Salle: Open Court, 1995); P.F. Strawson, "Social Morality and Individual Ideal," *Philosophy* 36 (1961): 1-17; F.A. Hayek, *Rules and Order* (Chicago: University of Chicago Press, 1973).

³⁹ Baier, *The Rational and the Moral Order*, p. 218.

⁴⁰ Rawls, *Political Liberalism*, p. 387. Cf. *The Order of Public Reason*, chap. VII.

⁴¹ Rawls, *A Theory of Justice*, revised ed. (Cambridge, MA: Harvard University Press, 1991), p. 106.

⁴² It is, I believe, an error to read Mill's *On Liberty* as a narrowly perfectionist or utilitarian argument. Showing how the case for liberty can draw on diverse perspectives, I have argued, is basic to Mill's argumentative strategy. See my "Controversial Values and State Neutrality in *On*

Liberty" in Mill's *On Liberty: A Critical Guide*, ed. C.L. Ten (Cambridge: Cambridge University Press, 2008): pp. 83-104.

⁴³ Benjamin Constant, "The Liberty of the Ancients Compared with that of the Moderns" in *Political Writings of Benjamin Constant*, ed. B. Fontana (Cambridge: Cambridge University Press, 1988), pp. 308-28. I have omitted here rights concerning political representation.

⁴⁴ As I argue in *The Order of Public Reason*, chap. VI.

⁴⁵ *Ibid.*, chap. VII. The problem of equilibrium selection is a general and fundamental problem of normative behavior; for insights into how this problem may be solved, see Bicchieri, *The Grammar of Society*.

⁴⁶ Paul Dragos Aligica, "Heterogeneity and Governance: Revisiting an Ostromian Thesis." Paper Delivered to the 2011 meeting of the American Political Science Association.

⁴⁷ In his *Lectures on the History of Political Philosophy Rawls* tells us that "a normalization of interests attributed to the parties" is "common to social contract doctrines." *Lectures on the History of Political Philosophy*, ed. S. Freeman (Cambridge, MA: Harvard University Press, 2007), p. 226.

⁴⁸ For a more detailed argument on this point, see "A Tale of Two Sets: Public Reason in Equilibrium," *Public Affairs Quarterly* 25 (October 2011): pp. 305-25.

⁴⁹ Jonathan Quong, *Liberalism Without Perfection* (Oxford: Oxford University Press, 2011), p. 167.

⁵⁰ For a fascinating study of the place of this authority relation in our attribution of moral responsibility, and what may be left of our normative relations with those who cannot make sense of it, see David W. Shoemaker, "Psychopathy, Responsibility, and the Moral/Conventional Distinction," *The Southern Journal of Philosophy* 49, Special Supplemental (2011): pp. 99-124.

Political Philosophy

G. A. Cohen's Critique of Rawls's Difference Principle

By Samuel Freeman

THE DIFFERENCE PRINCIPLE IS ONE OF THE MORE ORIGINAL BUT LEAST understood aspects of John Rawls's theory of justice. The difference principle says that inequalities of income, wealth, and economic powers and positions are to be arranged so as to maximally benefit members of the least advantaged group in society. There are many ambiguities implicit in the various formulations of the difference principle,¹ and the principle has been widely criticized for its apparent implications. One of the most enduring criticisms of the difference principle, emphasized by many to the left of Rawls, is that it supports the vast inequalities that typify capitalist economies. In his extensive treatment of Rawls's theory of economic justice, G.A. Cohen (now sadly deceased) extensively develops the argument that Rawls's difference principle justifies the "severe inequalities" of capitalism.² Cohen is especially concerned with the inequalities and rewards that go to the relative few whose exceptional talents are a result of the "natural lottery." Cohen contends that inequalities are undeserved in so far as individuals are not responsible for them and/or they are the product of chance. The natural talents that people are born with are undeserved, just like the social class they are born into and accidents and misfortunes they suffer during their lives. From this Cohen argues that differences in relevant goods (such as income and wealth, or individual welfare) that are the consequence of undeserved inequalities are also undeserved; therefore, he concludes, income and wealth should be redistributed to those who are naturally or socially less advantaged so as to equalize social advantages or to compensate them for arbitrary natural inequalities and other accidental contingencies. Cohen says that his position—termed (rather infelicitously) "the anti-arbitrariness conception of justice" by Cohen and "luck egalitarianism" by Elizabeth Anderson—is one that Rawls himself endorses early in *A Theory of Justice*, only later to abandon it with the difference principle. For the difference principle allows for incentives that encourage people to exercise their talents, and thus allows them to profit from arbitrary natural

Samuel Freeman is the Avalon Professor of the Humanities at the University of Pennsylvania. His research interests focus on social and political philosophy, ethics, and philosophy of law. He has written books on Justice and the Social Contract (Oxford 2006), and on the political philosophy of John Rawls (Rawls, Routledge, 2007), who was his dissertation supervisor. He has also edited the Cambridge Companion to Rawls (2003), as well as John Rawls's Lectures on the History of Political Philosophy (2007) and his Collected Papers (1999). He is currently working on a manuscript entitled Liberalism and Distributive Justice.