

## Introduction

---

*Paul Weithman*

The year 2021 saw the fiftieth anniversary of the publication of John Rawls's *A Theory of Justice*, a work that was said to revive political philosophy when it was published in 1971. To mark the anniversary, and the centenary of Rawls's birth, the University of Notre Dame hosted a two-day international conference: John Rawls's *A Theory of Justice* at Fifty. Planning for the conference began with the issuance of invitations in 2018. While the conference itself was held in the midst of a global pandemic that could not have been foreseen when planning began, it took place in the window between the spread of the Delta and Omicron variants of the coronavirus. Some who hoped to attend were unable to do so because of public health protocols. But the conference drew 140 registrants from fifteen states and a half dozen countries. For almost everyone in attendance, the conference was their first since the onset of the pandemic. Their shared sense that the academic world was opening up again made the occasion all the more celebratory.

Because planning for the conference began so far in advance, it was possible to secure commitments from many of the best political philosophers now working. The conversation within and between sessions was deep and probing. While it is impossible for a volume of conference essays to recreate the excitement of a conference, it is to be hoped that this volume will give readers some sense of the very high quality of work done at the conference and will advance both Rawls scholarship and political philosophy.

It is unfortunate that the volume does not include essays by two philosophers who were to have taken part: Jerry Gaus and Charles Mills. Their deaths are great losses to the field and their absence from the conference was sorely felt. But Jerry's position was ably represented by his student and friend Kevin Vallier, whose essay appears here. Charles Mills's illness would have prevented him from traveling to Notre Dame, but he had hoped to deliver his remarks by zoom. Sadly, he passed away days before the conference and we missed the benefit of even his virtual presence. Tommie Shelby and Henry Richardson, who

2 *Paul Weithman*

were to have shared the podium with him, paid tribute to him in the session in which all three of them were to have spoken, as did the session chair James Sterba.

In the course of developing and revising his theory, Rawls contributed to – and sometimes set the agenda for – a number of sub-areas of political and moral philosophy. He was also an accomplished historian of philosophy whose readings of his predecessors have proven immensely influential. Because Rawls worked on so many questions and had such wide influence, it is understandable that there were no lines of inquiry that ran through the conference in its entirety and that unite all of the essays in this volume. There were, however, threads that bound together groups of essays, sometimes in unexpected ways. The essays in this volume have been grouped so as to reflect those thematic continuities, but there are also interesting intersectional dialogues among the essays that will be noted along the way.

### **I.1 Rawls and History**

Rawls lectured regularly on the history of political and moral philosophy. Many of his students have gone on to do important and creative historical work. It was therefore natural to have a section on Rawls and the history of philosophy at the conference and to devote Part I of the conference volume to it.

In Chapter 1, S. A. Lloyd argues that Rawls's theory provides what her title refers to as "taillight illumination." By that she means that Rawls's theory casts light backwards in the history of political philosophy, illuminating previously neglected features of his predecessors' work. The predecessor whom she takes Rawls to illuminate is Thomas Hobbes, on whom Lloyd has done groundbreaking work, beginning with Lloyd (1992). Lloyd argues that in light of Rawls's development of political liberalism, certain features of Hobbes's work can be seen in sharper relief. The feature on which she focuses is Hobbes's political conception of the person. Seeing that conception at work in Hobbes, she thinks, both raises the possibility that Hobbes pioneered political liberalism and helps to break the hold of the game-theoretic interpretations of Hobbes that have been so prominent in the literature on him.

Daniel Brudney's "The *Theory* Rawls, the 1844 Marx, and the Market" (Chapter 2) juxtaposes the Rawls of 1971 with a thinker with whom he is not often compared: the Marx of 1844. If Lloyd uses Rawls to provide taillight illumination of his predecessors, Brudney uses Marx to provide what we might call "headlight illumination" of Rawls. For he uses Marx to highlight an important but underappreciated element of

*A Theory of Justice* and to bring to light an unappreciated and worrisome tension in that work. More specifically, Brudney argues that Marx's true communist society would be characterized by a "give/receive" relationship in which citizens care for one another's well-being. Having isolated the defining features of that relationship, Brudney then argues that such a relationship would also characterize the well-ordered society of *A Theory of Justice*. That is the underappreciated element of Rawls's theory and the one that gives rise to the unappreciated tension.

Brudney argues that the market activity Rawls allows in a well-ordered society would threaten the relationships of mutual care to which he is committed. But Brudney also argues that restricting the role of markets so as to preserve those relationships may well run afoul of what members of a market society may take as a tenet of common sense, since in a market society citizens may believe that indifference rather than mutual care is natural to us. In that case, the market restrictions necessary to preserve the give/receive relationship among citizens might not be publicly justifiable by appeal to the common sense of citizens and the deliverances of social science, as Rawls's commitment to public reason requires. It seems that Rawls can have mutual care or mutual justifiability but not both. In raising the possibility that mutual care cannot be had, Brudney's is the first essay in this collection – but not the last – to raise the possibility that the society Rawls envisioned may be vulnerable to, rather than stabilized by, the internal dynamics its institutions generate.

Aaron James's "Rawls, Lerner, and the Tax-and-Spend Booby Trap: What Happened to Monetary Policy?" (Chapter 3) locates Rawls, not in the history of philosophy, but in the history of economic thought – specifically in the economic thought of the mid-twentieth century. James argues that Rawls, like much of political philosophy, accepts the axioms of what he calls "sound finance." According to sound finance, budgets should be balanced and expenditures – including transfers and social programs needed to satisfy the principles of justice – should be paid for out of tax revenues. Accepting the second of these axioms leaves Rawls vulnerable to Nozick's famous critique of redistributive views: that the taxation needed to finance them is tantamount to forced labor and is therefore morally unacceptable.

James notes that as early as the 1930s and '40s some economists – notably Abba Lerner, whom James says Rawls read with care – had developed an alternative to sound finance: "functional finance." These economists argued that the view of taxation as a revenue-raising device is obsolete in a society that has moved off the gold standard. Once the gold standard has been abandoned, central banks can provide all the resources societies need – including all that they need to satisfy the difference

principle – by increasing the money supply. The question James raises is why Rawls ignored functional finance.

This is not just a historical question, since the embrace of sound finance left Rawls open to Nozick's objection and since functional finance has recently been championed in an important new work of economics (Marglin 2021). Moreover, James observes in passing that Rawls's acceptance of sound finance led him to rely on markets, suitably regulated, to achieve distributive justice. If Brudney is right about the tension in Rawls's view induced by his reliance on markets, then the question James raises about Rawls's neglect of functional finance is even more pressing than James says.

In "Rawls's Principles of Justice as a Transcendence of Class Warfare" (Chapter 4), Elizabeth Anderson locates Rawls in the history of social democratic thought. She begins by observing that in the 1980s, corporate elites in the United States initiated a series of assaults on the New Deal consensus. They justified their largely successful assaults and their economic gains by meritocratic arguments. Those arguments, Anderson says, are similar to the desert-based arguments wielded by elites in the class warfare of the late nineteenth century. But, she observes, it was not only nineteenth-century elites who appealed to desert. Representatives of all different class interests appealed to desert – albeit different conceptions of desert – to ground their claims. Where all such arguments misfire, Anderson says, is in the confusion of local and systemic principles. Local principles "directly guide agents in distributing goods to specific individuals." Systemic principles "govern the choice of *systems* of local principles by constraining their overall distributive consequences, defined in terms of opportunities for various goods." Desert may be appropriate for local application, but it is an inappropriate basis for systemic principles of distributive justice.

One side may gain the upper hand in class warfare if it successfully enforces its desert claims. But Anderson argues that class warfare can be transcended only with general acceptance of appropriate systemic principles grounded on shared democratic citizenship. This approach was pioneered early in the twentieth century by the Social Democratic Party of Sweden. "Rawls's principles of justice," Anderson says, are also grounded on common citizenship and so "aim to ... end[] a class-based society" and thereby transcend class warfare. She warns that the elimination of class distinctions may not lead to the elimination of other forms of hierarchy. But by attempting to transcend class warfare, and to envision a society in which all share the status of free equals, Anderson thinks Rawlsians fall squarely within the tradition of social democracy.

Peter de Marneffe's essay "The Significance of Injustice" (Chapter 5) is perhaps the most provocative in the collection. De Marneffe begins

with what he calls “the traditional view of justice.” That is a view of justice he finds in Plato, Aristotle, Aquinas, Hobbes, Locke, Adam Smith, Hume, and Kant. According to that view, injustice entails that a person suffering an injustice has been wronged. “Injustice in the traditional sense,” de Marneffe says, “has a special kind of significance for our relations to each other.” But, he continues “the failure of our social institutions to conform to Rawls’s two principles of justice does not have [that significance].” The reason it does not, de Marneffe argues, is that institutions can fail to conform with Rawls’s principles without anyone having performed actions that would warrant the reactive attitudes the tradition thinks are natural responses to injustice. It follows that Rawls’s claim to the contrary notwithstanding (*TJ* 1999, 9–10), his theory of justice does not “tally with” the traditional view of justice.

De Marneffe concludes that what Rawls has given us is not a theory of right and wrong or part of a theory of right and wrong. What he has given us is “a vision of the ideal democratic society” – perhaps, though de Marneffe does not say so, a vision of an ideal social democratic society à la Anderson. That society may be “something worthy of aspiration” but, de Marneffe concludes, realizing that society is “not necessary for mutual respect.”

## **I.2      Developments between *A Theory of Justice* and *Political Liberalism***

Part II of the volume is made up of four essays on the development of Rawls’s views in the crucial two decades between the publication of *A Theory of Justice* and the publication of *Political Liberalism*. It opens with an essay – Stephen Darwall’s “On Being a Self-Originating Source of Valid ‘Claims’” (Chapter 6) – that dovetails nicely with de Marneffe’s.

Darwall notes that Rawls introduces his description of persons as “self-originating sources of valid claims” in his Dewey Lectures, published in 1980. The Dewey Lectures are the writings in which Rawls’s Kantianism was at its apogee. Darwall argues that by adding the description of persons as self-originating sources of claims to his theory, Rawls filled a lacuna or remedied a defect in the Kantian interpretation of the theory he had laid out in section 40 of *A Theory of Justice*.

In section 40, Rawls says that we express our nature when we act from principles chosen in the original position. A society in which everyone realizes her nature may be what de Marneffe called “something worthy of aspiration.” But, Darwall argues, the fact that we express our nature when we act from the principles does not imply that those principles are principles of right. For Darwall, like de Marneffe, thinks that

principles of right are principles whose violation properly elicits certain reactive attitudes, such as guilt from oneself and blame from others. A failure to express our nature – or, as de Marneffe might have it, to realize something worthy of aspiration – elicits attitudes of an entirely different kind. Once persons are conceived of as self-originating sources of valid claims, however, they are conceived of as having the authority to hold one another accountable, and so to blame one another, for their failures. Parties in the original position, who represent persons so conceived, then choose principles that can properly be called principles of right.

Darwall does not address the question of whether violations of Rawls's two principles, which apply to the basic structure, properly occasion reactive attitudes with respect to other persons. His essay does not, therefore, have implications for the most fundamental worry de Marneffe raised. But it does show how Rawls supplemented his view after the publication of *Theory* so as to incorporate what de Marneffe calls "the traditional view of justice." The essay also shows, in a very satisfying way, how Darwall's own pioneering work on the second-person standpoint bears on the interpretation and viability of Rawls's theory.

Samuel Scheffler's "Moral Independence Revisited" (Chapter 7) also concerns the development of Rawls's thought after publication of *A Theory of Justice* – in Scheffler's case, the development that eventually led Rawls to recast his theory as a political liberalism.

In the introduction to *Political Liberalism*, Rawls credits Scheffler (1979) with having played an important role in that development. Scheffler begins his contribution to this volume by reviewing the argument of his essay and by asking what it was about that essay that led to such significant changes in Rawls's thought. To answer that question, he turns not only to Rawls's published work – prominently including "The Independence of Moral Theory" and the Dewey Lectures – but also to unpublished correspondence and to unpublished lectures that Rawls delivered in the late 1970s and 1980s. The result is an informed and nuanced discussion of Rawls's attempt simultaneously to establish political philosophy's independence from metaphysical questions about personal identity and philosophy of mind, and to give his conception of the person the central place that constructivism demands. Scheffler concludes by observing that Rawls thought he finally resolved the tension between those goals only in *Political Liberalism*, where he developed the political conception of the person and the political version of constructivism.

Rainer Forst's "The Method of Insulation" (Chapter 8) – like Darwall's and Scheffler's essays – treats of the development of Rawls's

thought between *A Theory of Justice* and *Political Liberalism*. His contribution, like Scheffler's, attends closely to Rawls's "Independence of Moral Theory" and his Dewey Lectures. Like Scheffler, Forst zeroes in on Rawls's attempt to insulate political philosophy from other areas of philosophy. And like Scheffler, Forst discerns a tension in Rawls's view, albeit a different one than Scheffler identified. The tension Forst discerns is between (i) the autonomy or insularity of a political conception of justice from what Rawls famously called "comprehensive doctrines" and (ii) the political conception's dependence on those doctrines for the moral force they need to trump unreasonable views. Forst is far less sanguine than Scheffler about Rawls's ability to resolve the tension he discerns. He concludes that "the struggle for ever more independence led to the danger of philosophical and normative dependence, and the island that Rawls tried to create is in danger of being washed away."

Japa Pallikkathayil opens "The Stability or Fragility of Justice" (Chapter 9) by laying out the concerns about *Theory's* treatment of stability that, Rawls says, led him to recast justice as fairness as a political liberalism. That Rawls should have been moved by these concerns shows that he regarded stability as an important feature of a conception of justice. Pallikkathayil asks why he accords it such importance, a question she pursues by addressing Gerald Cohen's claim that stability is alien to justice (Cohen 2008, 328). She argues that Rawls takes stability to be an important feature of justice as fairness because he thinks the role of that conception is to establish fair terms of cooperation and that any such terms are bound to be stable.

But she also contends, in sympathy with another of Cohen's criticisms, that the concept of justice has application even in circumstances where mutually advantageous cooperation is not possible – hence not only in what Rawls identifies as the circumstances of justice. "What matters" for the applicability of justice, she says, is precisely the claim that Darwall said Rawls needed to make if the parties in the original position are to choose principles of justice properly so called: "that we are 'self-authenticating sources of valid claims.'" Pallikkathayil insists that broadening the circumstances of justice beyond the circumstances of cooperation does not require giving up what she regards as the core Rawlsian insight: that the point of justice is to establish an inherently stable relationship of mutual justifiability among citizens. But she concludes by suggesting that a relationship of mutual justifiability might not be secured by the Rawlsian ideal of public reason. Stability and mutual justifiability may be achievable, she thinks, but they may require a non-Rawlsian conception of how citizens reason together.

### **I.3 Rawls, Ideal Theory, and the Persistence of Injustice**

Essays in Part III of the volume ask whether Rawls's theory is capable of handling cases of injustice, prominently including racial injustice. As would be expected, Charles Mills figures prominently in all of the essays here. It is regrettable beyond measure that he was not able to contribute an essay of his own to this part of the volume.

The first essay in Part III, like the final essay in Part II, argues for modifying Rawls's account of the circumstances of justice. In "The Circumstances of Justice" (Chapter 10), Erin Kelly argues that the circumstances should be taken to include contingent facts about a society's history of injustice and exploitation. Kelly is especially interested in the US history of racial injustice. Such historical circumstances make justice necessary and, Kelly argues, parties in Rawls's original position should take account of them when adopting principles of justice. Thus does she think the resources of Rawlsian political philosophy can be brought to bear on legacies of domination and structural injustice. Moreover, Kelly thinks that once we see Rawlsian principles as remedies to historic injustice, we can add to the reasons Rawls offers for thinking that justice is a good. Justice is a good because it helps to effect moral repair in response to collective wrongdoing. Kelly contrasts her approach with the way it is usually thought Rawls would handle problems of historical injustice, by relegating them to the realm of non-ideal theory. She concludes by sketching some affinities between her approach to the grave injustice of racism and the approach of Charles Mills.

Henry Richardson begins "Why Rawls's Ideal Theory Leaves the Well-Ordered Society Vulnerable to Structural Oppression" (Chapter 11) with the frank acknowledgment that Mills's work reveals a flaw in Rawlsian ideal theory. The well-ordered society of justice as fairness does not contain adequate bulwarks against racism, so even a just society could eventually come to be marred by bigoted oppression. Richardson's conclusion recalls and promises to vindicate Elizabeth Anderson's closing warning that elimination of economic class hierarchy does not mean the elimination of hierarchies of other kinds. And it anticipates arguments by Kevin Vallier and Joshua Cohen that the justice of Rawls's well-ordered society may be more vulnerable than Rawls acknowledged.

The vulnerability of a well-ordered society is due, Richardson argues, to Rawls's overly juridical understanding of the basic structure and his overly moralized conception of power. The basic structure as Rawls conceives it consists only of some legally constituted institutions and of the legally constrained aspects of others. Power as Rawls conceives it is

overly moralized because it excludes various kinds of de facto power and the power exercised by informally constituted in-groups. Members of these groups may gain power through access to cultural and social capital that others do not enjoy. Even in a well-ordered society, the elite snobism that can develop among those with such access can eventually – through stages Richardson hypothesizes – lead to exclusion and domination that are unjustifiable but beyond the reach of law. But while Charles Mills criticized Rawls for pursuing ideal theory, Richardson believes the flaws he identifies in Rawlsian ideal theory can be repaired; how to repair them is the one of the subjects of Richardson's ongoing work.

In “Race, Reparations and Justice as Fairness” (Chapter 12), Tommie Shelby acknowledges that Rawlsian ideal theory would be gravely flawed if it could not condemn racial injustice. He has argued in other work that it can. In this essay, he argues – contra Charles Mills – that reparations for past racial injustice are consistent with the essentials of Rawls's theory. Thus he argues that parties in the original position would acknowledge a natural duty of reparation. He also argues, in the spirit of the priority Rawls accords the basic structure as a subject of justice, that attaining a just basic structure should have priority over the duty to make reparations for past injustices and that reparations made for recent injustices must be consistent with attaining and maintaining a just basic structure.

In “On the Role of the Original Position in Rawls's Theory” (Chapter 13), Laura Valentini asks what a normative theory is and insists on the importance of distinguishing the desiderata of normative theories from the evidence supporting them. This distinction is, she says, crucial to rebutting two quite different critiques of Rawls. One is Charles Mills's critique that Rawls's theory is excessively idealized. The other is a criticism touched on by Pallikkathayil's essay: Gerald Cohen's critique of Rawls's constructivism as fact-sensitive.

Both Mill's and Cohen's criticisms of Rawls zero in on the role of the original position. But, Valentini argues, these criticisms are not criticisms of Rawls's theory properly speaking because the original position is not, properly speaking, part of Rawls's theory. Arguments from the original position are evidence for, rather than part of, that theory. Rawls's aims justify his reliance on evidence provided by an idealized choice situation in which principles are adopted for idealized conditions. Mill's critique of ideal theory therefore misfires. Moreover, given those aims – which include formulating an account of justice that can enduringly well-order human societies in the circumstances of justice – a theory that has those aims must be fact-sensitive. So Cohen's critique of Rawls's principles as

fact-sensitive misfires as well. Of course one can reject Rawls's aims, as Mills and Cohen seem to do. But, Valentini insists, to reject the aims of a theory is not to reject the theory itself.

#### **I.4 Pluralism, Democracy, and the Future of Justice as Fairness**

Kevin Vallier's "Public Reason at Fifty" (Chapter 14) concerns what he calls "the public reason project" as it developed through Rawls's writings and beyond. The public reason project is the project of determining what basic social arrangements, if any, can be justified to everyone who is subject to them. Rawls clearly pursued that project. Indeed, we saw that the argument of Daniel Brudney's essay turned on Brudney's claim that mutual justifiability by public reason is one of Rawls's central commitments. The late Gerald Gaus also pursued it. Vallier defends the superiority of Gaus's approach to the project because, he thinks, it effectively responds to the failure of Rawls's.

To see the failure that Vallier thinks undermines the Rawlsian approach, recall a point Pallikkathayil discussed in her essay: Rawls thinks the best conception of justice is the one that is most likely to be enduringly or stably adhered to once it is instituted. The Rawls of *A Theory of Justice* defended justice as fairness on the ground that it satisfied that desideratum. Vallier presents the development of Rawls's thought after *Theory* as a series of attempts to remedy what he found unsatisfactory in that defense. Vallier contends that the free society Rawls's principles of justice require is bound to be characterized by reasonable pluralism about justice. The three models of public reason Vallier says Rawls developed in *A Theory of Justice*, and then in *Political Liberalism* and its paperback edition, attempt first to deny and then to contain such justice pluralism.

Rawls's hope was to show that the desideratum he asserted in *Theory* would be satisfied by justice as fairness or by a small family of liberal conceptions of justice. But, Vallier thinks, none of the Rawlsian models succeeds. Gaus's response to this failure developed significantly over many years; Vallier identifies four models of public reason in Gaus's writings. Though Vallier does not say so here, it seems that the momentum of Gaus's thought was carrying him toward the denial that stability is a desideratum. If that is right, then a non-Rawlsian conception of public reasoning leads to a very different place than Pallikkathayil suggests at the end of her essay.

Samuel Freeman (Chapter 15), like Vallier, is interested in the ways Rawls revised his view to accommodate the possibility of what Vallier