The Tyranny–or the Democracy–of the Ideal?

BLAIN NEUFELD AND LORI WATSON

INTRODUCTION

Gerald Gaus's *The Tyranny of the Ideal* is an ambitious book that covers an impressive range of topics in political philosophy and the social sciences. The book launches a systematic critique of 'ideal theorizing' about political and social justice and aims to defend a vision of an 'Open Society' that "forsakes a collective ideal of justice" (Gaus 2016, p. xvi).[1] Gaus charges that the dominant philosophers of justice in recent decades, at least within the Ango-American tradition, have been seduced by the allure of identifying 'the' ideal conception of justice. Part of this approach to theorizing about justice involves positing "well-ordered societies, where we [the citizens] all agree on the correct principles of justice, our institutions conform to them, and we are all committed to them" (Gaus 2016, p. xix). Ideal theorizing, so construed, is understood by many political philosophers to provide a useful guide for reforming and reshaping our present non-ideal and unjust societies. Gaus argues, though, that this kind of ideal theorizing about justice "tyrannizes over our thinking, preventing us from discovering more just social conditions" (ibid.). The book aspires, then, to show political philosophers that many of them have been labouring under a yoke that they have failed to recognize. More than this, the book also aims to articulate an alternative approach to ideal theorizing, one that frees theorists from this tyranny.[2]

Unsurprisingly, John Rawls's political philosophy is a central focus of Gaus's 'liberation project.' Rawls's work, including its form of ideal theorizing,[3] has significantly shaped the field of political philosophy since the publication of *A Theory of Justice* in 1971. Gaus argues, though, that Rawls's later attempt to accommodate the fact of reasonable pluralism in *Political Liberalism* (2005) leaves his overall theory of justice "in disarray" (Gaus 2016, p. 153). Further, Gaus contends that once we fully confront the depths of diverse points of view concerning justice, not only must Rawlsians abandon the ideal of the well-ordered society, they must reorient their thinking about ideal theory altogether.

In contrast to Rawls, Gaus holds that a morally heterogeneous society—a society characterized by deep differences concerning justice and not merely 'conceptions of the good' and 'comprehensive doctrines'[4]—is necessary for us to advance our knowledge of the requirements of justice. Abandoning the 'myth' of the well-ordered society and replacing it with the idea of the Open Society, including especially its embrace of diverse and evolving perspectives on justice, is much more likely to "make the world better for all, and allows us to better understand our different moral truths" (Gaus 2016, p. xx). Thus instead of trying to cultivate the one 'perfect rose,' so to speak, Gaus recommends that political philosophers encourage 'a thousand flowers to bloom.'

In this article we defend the Rawlsian view against Gaus's criticisms.[5] Specifically, we dispute Gaus's claim that Rawls's idea of the well-ordered society cannot survive the move to political liberalism. We formulate a 'political liberal' version of the well-ordered society, and show that Gaus's Open Society, rather than a radical alternative to the political liberal well-ordered society, in fact closely resembles it. We also challenge Gaus's claim that Rawlsians committed to the principles of justice as fairness are confronted with 'The Choice.' According to The Choice, roughly, ideal theorists must either: (1) pursue 'nearby' relatively certain 'local' gains in justice for their society, or (2) forgo these local gains in order to pursue the more ambitious but far less certain goal of 'ideal' justice. (The goal of ideal justice is 'less certain' both in terms of its likely achievement as well as the likelihood that it is in fact *the* ideal.) We challenge Gaus's claim regarding The Choice, at least as applied to the Rawlsian view, by explaining how addressing local injustices naturally can *lead* some citizens to develop conceptions of full justice, including 'realistically utopian' versions of their societies. The kinds of political proposals that plausibly follow from this account of public reasoning indicate that Rawlsians in fact do not confront The Choice (or, at the very least, that some additional argument is needed to show that they do). Thus, despite the many interesting points that Gaus raises in his book, we conclude that his arguments do

not ultimately threaten the Rawlsian approach to thinking about political justice.

I. TWO PUBLIC REASON LIBERALISMS

Before we get to our replies to Gaus's criticisms of Rawlsian ideal theorizing (in Sec II-III), we think that it would be helpful to make some general points about Gaus's and Rawls's respective projects in political philosophy. As is well known, Rawls resurrected philosophical interest in political contractualism with the publication of *A Theory of Justice*. In that work, Rawls employs the device of the 'original position' (See Rawls 1999, Sec 4 and Ch. III)—a device that, in Gausian terminology, 'normalizes' the perspectives of diverse citizens (Gaus 2016, pp. 42-51, 105-114) by (inter alia) placing the parties who represent those citizens behind a 'veil of ignorance,' thereby depriving them of any particular knowledge concerning the identities of the citizens whom they represent. The parties within the original position consequently all reason in the same way and have access to the same (general) information. Rawls proposes in *Theory* that the parties would select the two principles of 'justice as fairness,'[6] and that a society 'well ordered' by those principles would be stable over time. Such a society would be stable because all of the citizens who belong to it would endorse and support over time—via their developed and rationally maintained 'sense of justice' (Rawls 2001, pp. 18-19)—the institutions of their just 'basic structure.'[7]

Eventually Rawls came to have doubts about *Theory*'s account of the stability of the well-ordered society.[8] Simplifying greatly, this is because part of that account rested on the acceptance by all citizens of a broadly Kantian ideal of autonomy. Rawls came to think that not *all* citizens within the well-ordered society could endorse this ideal. Instead, the citizens of any society that conformed in its basic structure to the principles of justice as fairness, through the free exercise of their reason, invariably would come to endorse a plurality of reasonable 'comprehensive doctrines' (religious, moral, and philosophical views). This 'fact of reasonable pluralism' (Rawls 2005, p. 441) motivated Rawls to develop his theory of political liberalism, of which the idea of 'public reason' is a central component.

In his writings on political liberalism Rawls holds that decisions concerning fundamental political questions—those having to do with "constitutional essentials" and "matters of basic justice" (Rawls 2005, pp. 214-15, 227-30, 235)—should be made by means of *shareable* public reasons (reasons that all reasonable citizens find acceptable, despite their adherence to different comprehensive doctrines). The idea of public reason, Rawls proposes, should be understood as "part of the idea of democracy itself" (Rawls 2005, p. 441). This is because by deciding fundamental political questions via shareable public reasons, or by ensuring that their political representatives do so, citizens can relate to one another as equal co-*sovereigns*. Citizens also are the *subjects* of political decisions. Political power is ultimately coercive in nature,[9] but the exercise of such power over citizens can be normatively legitimate—it can satisfy what Rawls calls the 'liberal principle of legitimacy'—if it is authorized by a constitutional structure that is justified in terms that all citizens find acceptable (Rawls 2005, pp. 37, 446-447).

Public reasons are directed at a moderately idealized justificatory constituency: citizens whom Rawls labels 'reasonable persons.' Reasonable persons acknowledge the fact of reasonable pluralism and are committed to what Rawls call the "criterion of reciprocity" (Rawls 2005, pp. 48-58). According to this criterion, decisions regarding constitutional essentials and matters of basic justice must be *acceptable* to those citizens subject to them (even if those decisions are not the most preferred ones of all citizens). Satisfying the criterion of reciprocity, then, involves citizens providing mutually acceptable justifications for their shared exercise of political power.[10]

Among Rawlsian public reasons are ecumenical democratic ideals and civic virtues, like transparency and toleration, as well as general rules of inquiry, such as those concerning evidence, logic, and so forth. Public reasons also can be drawn from 'reasonable political conceptions of justice.' Conceptions of justice are 'reasonable' *if* they satisfy the criterion of reciprocity.[11] Such conceptions, in order to satisfy this criterion, must secure a set of specially ranked 'basic liberties' for all citizens (including liberty of conscience, freedom of association, and the political liberties of democratic citizenship), as well as adequate resources (such as education and wealth) for all citizens to exercise effectively those liberties over the course of their lives (Rawls 2005, p. 450). A reasonable conception of justice is 'political' if it is compatible with the various comprehensive doctrines endorsed by reasonable citizens, that is, if it is 'freestanding' in nature. A political conception of justice also is limited in its scope to the basic structure of society;[12] hence it does not apply to all domains of social life. ('Comprehensive' conceptions of justice, in contrast, presuppose the truth of particular comprehensive doctrines, such as utilitarianism, and/or apply directly to domains of social life beyond the basic structure.)

A reasonable political conception of justice—because it satisfies the criterion of reciprocity, is compatible with citizens' different comprehensive doctrines (is freestanding), and is limited in its scope to the basic structure—can be the focus of what Rawls calls an "overlapping consensus" (Rawls 2005, Lecture IV). In such a consensus, roughly, citizens who endorse different comprehensive doctrines *also* can support the reasonable political conception of justice. They may do so by integrating the political conception of justice into their broader sets of beliefs and values.[13] A society in which there is an overlapping consensus on a reasonable political conception of justice consequently can be *stable* over time through the free allegiance of its members. Hence Rawls claims that a well-ordered society characterized by the fact of reasonable pluralism is possible. But, as Gaus stresses in *Tyranny*, in his later writings Rawls *also* acknowledges that there is a *family* of reasonable political conceptions of justice. It is not clear whether the idea of a well-ordered society based on an overlapping consensus coheres with Rawls's acknowledgement of a plurality of reasonable political conceptions of justice. (We try to explain why these ideas are compatible, albeit with some modifications to the idea of a well-ordered society, in Sec II.)

In recent years Gaus has developed an alternative 'convergence' account of public reason justification[14] to Rawls's 'consensus' account. With respect to legislation in general, and not simply constitutional essentials and matters of basic justice (the domain, recall, to which Rawlsian public reasons primarily apply),[15] Gaus applies what he calls the "Public Justification Principle." This principle states: "*L* is a justified coercive law only if each and every member of the public *P* has conclusive reason(s) *R* to accept *L* as binding on all" (Gaus 2010, p. 244). Hence Gaus holds that a political decision can be legitimate if all relevant citizens—the moderately idealized justificatory constituency of the "members of the public" (Gaus 2011, p. 26)—have (at least) *a* sufficient reason to support it. Different members of the public, however, can rely upon different, even incompatible, reasons to support laws. For instance, some citizens might use reasons drawn from their respective religious doctrines while others might appeal to philosophical views like utilitarianism. It is for this reason that Gaus's account of public justification often is referred to as a 'convergence' account: diverse justifications can 'converge' in supporting a law, and thereby secure the legitimacy of that law, even if there is no 'consensus' amongst all members of the public on any of those justifications.

Gaus thus denies a central claim that Rawls advances in *Political Liberalism*: namely, that a political conception of justice requires a *freestanding* justification, and that that freestanding justification must be 'political' in nature (that is, draw upon shareable, ecumenical political ideas, like the ideals of citizens as free and equal, and society as a fair system of social cooperation) rather than 'comprehensive.'[16] While Gaus does not explicitly refer to the Public Justification Principle in *Tyranny*, he does make use of the same justificatory structure with respect to the relation between citizens' diverse conceptions of, or perspectives on, ideal justice, and the set of social rules (including, but not limited to, coercively-enforced laws) that apply to all, and to which all consequently are subject and mutually-accountable.

Gaus's 'Open Society' is regulated by what he calls a 'public moral constitution.' Simplifying somewhat, this public moral constitution consists of a set of rules (that evolve over time) for social cooperation, and with respect to which persons hold one another accountable (Gaus 2016, pp. 177-240). Thus we can say that there is a kind of 'overlapping consensus' within the Open Society; however, the public moral constitution need not have its own *independent* justification (a justification that draws on shared normative ideas) that serves as the basis for public reasoning (Gaus 2016, pp. 177-78). Rather, the Open Society arises *out* of its members' practice of public reasoning together (Gaus 2016, p. 179), in which they draw upon their respective non-shared values, beliefs, and perspectives on ideal justice in order to arrive at mutually acceptable rules and practices.

In developing his account of the Open Society, Gaus recognizes that some 'partial normalization' of perspectives is required—namely, some normative assumptions about our co-operators' attitudes, beliefs, and dispositions. These features include the fact that the members of the Open Society wish to engage in a cooperative scheme, endorse norms of mutual accountability, see certain reactive attitudes as justified, and—most importantly for our discussion—do not take up what Gaus calls an "optimizing stance" (Gaus 2016, p. 215), that is, they do not insist that their ideal "be instituted in the face of disagreement by other perspectives" (Gaus 2016, p. 218) no matter what. One might restate this as: the co-operators in the Open Society are not dominators. Or, to put the point in Rawlsian terms, they desire a social world in which they can cooperate with others as free and equal on terms acceptable to them all.

So not all persons or all perspectives on ideal justice can belong to the justificatory constituency for the public moral constitution. Gaus (2016, p. 222) writes:

> Some perspectives are, in the end, unable to share a framework of moral accountability with diverse others. Even the Open Society must be prepared to normalize to some, hopefully very small extent. Such 'Excluded Perspectives,' which cannot find sufficient space in the Open Society, will almost surely be those that are committed to the optimizing stance […]. The Excluded perspectives can live only by those [rules] that they think best, and so cannot endorse the characteristic institutions of the open society.

So like Rawls's reasonable political conception of justice, Gaus's public moral constitution cannot be supported by *all* of the members of society. Despite their differences, then, both Rawls and Gaus appeal to *restricted* justificatory constituencies: reasonable persons in the case of Rawls, perspectives on justice not committed to the optimizing stance in the case of Gaus. Simply put, both approaches restrict their justifications to constituencies committed to a notion of reciprocity.

II. THE WELL-ORDERED SOCIETY AND THE OPEN SOCIETY

After laying out the arguments against Rawlsian ideal theory and presenting his own positive view of the role of ideal theorizing within the Open Society, at the end of the book Gaus concludes that we must bid "adieu" to the idea of the well-ordered society as a guiding ideal. This ideal is not a feasible or attractive goal; it can give us no useful practical guidance about what to do here and now, nor can it "help reconcile us to this conflict-ridden and often manifestly unjust social world" (Gaus 2016, p. 245). While different communities of citizens or research groups within the Open Society should be allowed or even encouraged to develop their respective conceptions of ideal justice—and to engage in constructive criticism and debate with each other over those conceptions—the hope that citizens might someday come to all endorse the *same* conception of ideal justice must be abandoned.

In the concluding chapter of *Tyranny*, Gaus summarizes two main charges against the ideal of the well-ordered society:

1. It is a mirage.
2. As such, it "tyrannizes over our thinking and encourages us to turn our backs on pressing problems of justice in our own neighbourhood" (Gaus 2016, p. 246).

With respect to the first charge, the ideal of a society in which all (reasonable) citizens accept the same conception of justice and know that all others accept it is a mirage, argues Gaus, because even *if* we could arrive at complete agreement about the correct principles of justice, we would not agree about which "social states best satisfied them" (Gaus 2016, p. 246). Moreover, it is hopelessly unrealistic to even think that we *could* settle upon one ideal of justice. One reason for this is that, as we get 'closer' to an ideal on which we may have agreed upon earlier, our knowledge concerning that ideal expands such that a *new* and improved ideal almost certainly will emerge.

With respect to the second charge, Gaus holds that the ideal of the well-ordered society can act as a will-o'-the-wisp, leading theorists to futilely pursue it and thereby ignore the immediate injustices of their society. This is because political philosophers' theorizing about ideal justice inevitably is subject to what Gaus calls the "Neighbourhood Constraint": "we have far better information about the realization of justice in our own neighbourhood than in far-flung social worlds" (Gaus 2016, p. 102).[17] That is, we can apprehend with relatively high confidence what a 'nearby' social world produced by modest institutional changes would look like; in contrast, we cannot be especially confident in our judgements about more 'distant' social worlds, ones produced by more radical or extensive institutional changes. And given the limitations of our knowledge, powers of prediction, accurate modeling, and so on—limitations reflected in the Neighbourhood Constraint—ideal theorists consequently face "The Choice": "In cases where there is a clear optimum within our neighbourhood that requires movement away from our understanding of the ideal, we often must choose between relatively certain (perhaps large) local improvements in justice and pursuit of a considerably less certain ideal, which would yield optimal justice" (Gaus 2016, pp. 82, 246). The ideal of the well-ordered society thus can tempt us to make 'the perfect the enemy of the good,' so to speak.

In order to mitigate (though likely not eliminate entirely) The Choice faced by the advocates of any particular ideal of justice, they must recognize that they will (almost certainly) improve *their* understanding of their own ideal if they have access to *other*, rival ideals of justice—and, moreover, can

expose their ideal of justice to criticisms, challenges, and recommendations for reform from others who do not share it. This epistemic gain from diversity, though, comes with a cost: a society that fosters a diversity of perspectives on ideal justice will (almost certainly) never converge on a *single* one of those perspectives. (Some conceptions of justice may come to be abandoned over time, of course, but others will be created, others will evolve, and so forth.) As Gaus puts this point: "Only those in a morally heterogeneous society would have a reasonable hope of actually understanding what an ideal society would be like, but in such a society we will never be collectively devoted to any single ideal" (Gaus 2016, p. xix). Thus philosophers cannot have their 'ideal justice cake' and eat it too. Society nonetheless can benefit from ideal theorizing about justice—but only once the aims of such theorizing are appropriately *chastened*. The activity of widespread 'non-tyrannical' ideal theorizing, by many different communities of citizens with different perspectives on ideal justice, can help society progress towards greater overall (non-ideal) justice. Hence ideal theorizing need not be pointless or counterproductive—even if *no* individual perspective will ever turn out to be endorsed by all citizens in any given society.

The main complaint that Gaus levels against Rawls's views as formulated in *Theory*, then, is that the argument for the two principles of justice as fairness rests on the 'normalization' of perspectives, that is, the elimination of diversity among deliberators (via the original position device). Such normalization is neither realistic nor will it produce an accurate ideal of justice. The "deep dilemma" that a normalized approach to ideal justice faces is this: "the very normalization that defines the 'correct' perspective on justice cannot effectively identify its own ideal" (Gaus 2016, p. 150). Rather, as explained above, Gaus holds that advancing our knowledge of justice requires a deep diversity of perspectives, because any given perspective (that is, any particular approach to thinking about justice, such as the Rawlsian one) aiming to identify its own ideal may do so, but probably only will have arrived at a 'local optimum' because it is blind to certain relevant considerations. This is because: "It is other, different but related perspectives, that are most likely to see overlooked superior alternatives—ones that the original perspective can appreciate it has overlooked" (Gaus 2016, p. 135). To use the metaphor of mountain climbing,[18] from any given perspective you may identify the highest peak in *your* range, but you may fail to identify *higher* peaks further afield, let alone the highest peak of all.

To illuminate this claim, Gaus reminds readers of the various insights feminist perspectives have brought to bear on our thinking about justice over the past forty years (Gaus 2016, p. 134). Indeed, early feminist criticism of *Theory* led to important revisions in Rawls's formulation of the original position.[19] Sex was added explicitly to the list of features of persons about which deliberators in the original position were deprived of knowledge behind the veil of ignorance. The specification of the parties in the original position as 'heads of households' also was dropped. This is but one example of how a diversity of perspectives can help improve any particular approach to thinking about justice.

Recall Rawls's original formulation of the well-ordered society: "it is a society in which (1) everyone accepts and knows that others accept the same principles of justice, and (2) the basic social institutions generally satisfy and are generally known to satisfy these principles" (Rawls 1999, p. 4). In a society that is so ordered, citizens "acknowledge a common point of view from which their claims may be adjudicated"; moreover, a shared public sense of justice "establishes the bonds of civic friendship" (Rawls 1999, p. 5). The idea of the well-ordered society is importantly tied to defending justice as fairness as capable of securing stability for the right reasons in Part III of *Theory*. The representatives in the original position assess conceptions of justice according to their potential for securing stability: more stable principles of justice are preferable to less stable schemes.[20] Rawls aims to show that justice as fairness will be more stable than alternative conceptions of justice (or at least stable enough):[21] it generates its own support and is more in line with principles of moral psychology than alternative conceptions (Rawls 1999, Part Three).

In the move from *Theory* to *Political Liberalism* Rawls makes clear, as we noted earlier, that justice as fairness is in fact but *one* member of a family of reasonable political conceptions of justice. This is a consequence of taking the fact of reasonable pluralism seriously—such reasonable pluralism does not simply include the range of reasonable comprehensive doctrines and conceptions of the good, but also pluralism with respect to justice itself. Thus, according to the final version of Rawlsian political liberalism, there is a family of reasonable political conceptions of justice from which reasonable citizens can draw when making public reason arguments in favour or against proposals concerning fundamental political questions. It is in light of this development that Gaus concludes: "One has to be an especially devout disciple of Rawls not to conclude that by the close of his political liberalism project the theory of justice

was in disarray" (Gaus 2016, p. 153). This is because Gaus holds that the idea of the well-ordered society cannot survive in light of the developments in *Political Liberalism*.

Gaus proposes that a liberal Open Society—and *not* a well-ordered society—provides the best framework for securing the necessary terms of social cooperation for citizens (including communities of citizens committed to different ideals of justice) to coordinate effectively their actions and hold each other accountable. Such an Open Society contains a wide range of perspectives on justice (as well as religious and philosophical views, conceptions of the good, and so forth). The Open Society nonetheless possesses an overarching "public moral constitution"—"a stable, shared, moral, and political framework for living together" (Gaus 2016, p. 178)—which (most of) the perspectives on ideal justice find acceptable, and thus with which they can *comply* over time.

We think that Gaus is right about the version of the idea of the well-ordered society formulated in *Theory*. However, it remains an open question whether the idea of the well-ordered society can be revised or reconstructed in light of political liberals' recognition that there is no single political conception of justice that we can realistically expect full agreement upon. In other words, once political liberals acknowledge that there is a range of reasonable political conceptions of justice, they must either revise the idea of a well-ordered society or abandon it. Here we aim to do the former: we provide a reconstruction of the idea of the well-ordered society as it functions in political liberalism. For the sake of convenience, we will refer to this as the 'PL WOS' (political liberal well-ordered society).

Our account of the PL WOS does not rely on a problematic complete normalization of citizens' perspectives (such as that criticized by Gaus with respect to *Theory*'s formulation of justice as fairness). Reasonable persons hold a diversity of political views. While they all are committed to the criterion of reciprocity—this commitment (along with their acceptance of the fact of reasonable pluralism) is what makes them 'reasonable'—they do not all hold that the original position is the best way to satisfy this criterion when thinking about political justice. As such, the members of the PL WOS cannot be expected to endorse a single political conception of justice (such as justice as fairness). Though, of course, citizens can draw upon their favoured conceptions of justice when making public reason arguments, doing so does not commit them to an 'optimizing perspective' whereby they *insist* that their conception of justice alone be the basis of their society's laws. The idea of public reason, based upon the criterion of reciprocity, thus creates the conditions for the exchange of public reasons, a process by means of which citizen can arrive at mutually acceptable justifications for the institutions and policies of their shared basic structure.

Philosophers cannot know in advance of the exercise of public reasoning what the outcomes of citizens' deliberations will be. However, as mentioned in Sec I, in order to satisfy the criterion of reciprocity, and thereby be 'reasonable,' a political conception of justice must include certain features. Specifically, Rawls holds that all members of the family of reasonable political conceptions of justice will contain the following three features:

> First, a list of certain rights, liberties, and opportunities (such as those familiar from constitutional regimes);
> Second, an assignment of special priority to those rights, liberties, and opportunities, especially with respect to the claims of the general good and perfectionist values; and
> Third, measures ensuring for all citizens adequate all-purpose means to make effective use of their freedoms (Rawls 2005, p. 450).

Political conceptions of justice that include some version of these three features embody respect for all citizens as free and equal persons.[22]

While reasonable citizens typically will find most plausible, and hence endorse, only *one* reasonable political conception of justice,[23] they will judge *all* conceptions that satisfy the criterion of reciprocity to be 'acceptable.' A citizen finds a political conception of justice 'acceptable' insofar as she can appreciate the justification(s) for that conception, and—because that conception satisfies the criterion of reciprocity, has a justification that is freestanding in nature, and contains principles that apply only to the basic structure—can willingly abide by its institutional requirements should it be implemented democratically in her society's basic structure. This is so even if that citizen would prefer a *different* conception of justice to be realized in her basic structure, that is, even if she regards an alternative political conception to be 'more reasonable.'[24] What is important for our purposes here is the idea that for a reasonable person to find a conception of justice acceptable is (ceteris paribus) sufficient reason for that person to *comply* with the various institutions and laws justified by that conception (if they are realized via legitimate political procedures).

With the idea of a family of reasonable political conceptions of justice in hand, we can formulate the PL WOS—a

society regulated by a public conception of justice (Rawls (2001, p. 8)—as a society with the following features:

1. All citizens (as reasonable persons) endorse *a* reasonable political conception of justice (one member of the family of reasonable political conceptions of justice).
2. The basic structure is organized in compliance with (at least) one member of the family of reasonable political conceptions of justice.[25]
3. All citizens (reasonable persons) know (1) and (2) ((that is, the 'publicity condition' (Rawls 2005, pp.66-71) is satisfied)).
4. A public political culture obtains as characterized by a reasonable overlapping consensus and a shared commitment (among reasonable citizens) to public reason.

Such a PL WOS can be 'stable for the right reasons,' namely, through the *free* support of the reasonable citizens who belong to it. And it is 'realistically utopian' in that citizens can try to more closely realize it through their political activity, and in particular by deciding fundamental political questions via public reasons, *without* expecting that everyone must endorse the same political conception of justice.

Thus, contrary to Gaus's claims, the idea of the well-ordered society, a pluralist society that is stable for the right reasons, can survive the move to political liberalism. The PL WOS does not require that all citizens share the same conception of justice; nor does it require the use of a single normalizing perspective (such as that of the original position). The PL WOS relies on the criterion of reciprocity, the family of reasonable political conceptions of justice, and the idea of public reason. The core of the ideal of the well-ordered society is that of a system of fair social cooperation (across generations) amongst free and equal citizens; the PL WOS realizes this ideal, despite pluralism with respect to justice.

But what is the fate of Rawls's specific conception of justice in the PL WOS? Gaus observes, "In 1958's 'Justice as Fairness' a 'family' of distributive views was also justified, and Rawls saw this as a core *weakness* in a theory of justice," because that family could not provide "a determinate ranking of claims" (Gaus 2016, p. 152, our emphasis).[26] Are we transported back to 1958, so to speak, in the PL WOS? Our answer is 'yes and no.' The 'yes' part of our answer involves recognizing that, for society as a whole, decisions regarding the 'ranking of claims' concerning justice are to be made by *citizens* as co-sovereigns and public reasoners. Philosophy cannot identify a single conception of justice on which all rational and reasonable citizens must converge. At the same time, though, for particular citizens our answer is 'no.' This is because the conception of justice as fairness will remain (at least for the foreseeable future) a member of the family of reasonable political conceptions of justice that citizens can endorse and draw upon when providing public reason justifications for political proposals.

Even in his final writings on political liberalism Rawls maintains that justice as fairness is the *most* reasonable conception of justice (See Rawls 2005, pp. xlvi-xlvii, 450-51). Holding this view is compatible with acknowledging that other conceptions are reasonable, and thus also can serve as the bases for legitimate laws. Political philosophers occupy no privileged position in the PL WOS. "[T]here are no philosophical experts," Rawls writes, "Heaven forbid!" (Rawls 2005, p. 427). But this is not to say, of course, that philosophers cannot *contribute* to the public political culture of their society. "[C]itizens must, after all, have some ideas of right and justice in their thought and some basis for their reasoning," Rawls observes, "And students of philosophy take part in formulating these ideas, but always as citizens among others" (Rawls 2005, p. 427). When we judge a particular conception to be the *most* reasonable one, we presumably have reasons for doing so—reasons that we can share with others, in the hope of convincing them also to endorse that conception. The original position thus can continue to serve as a tool by means of which some citizens can explain *why* they are committed to the conception of justice as fairness, and thus endorse (and support with public reasons) the political proposals that they think follow from that conception. Granted, it is not the *only* way to satisfy the criterion of reciprocity (Rawls 2005, pp. xlviii-xlix), but we think that it remains *an* important way, one that many citizens continue to find compelling, and consequently can draw upon in their political deliberations and activities.

Having revised the idea of the well-ordered society so that it is compatible with political liberalism, and in a way that we think expresses fidelity to Rawls's main commitments, we find it difficult to see the stark contrast that Gaus aims to draw between it and the Open Society. The Open Society, like the PL WOS, is deeply concerned with securing the support and compliance of its members. With respect to the public moral constitution of the Open Society, Gaus (2016, p.178) writes:

> The aspiration is for the various perspectives, each committed to its understanding of the nature of the social world and ideal justice, to find the public social

> world *endorsable*. If each perspective can make sense of the categories of the artificial social world and *endorse* their use […], we can have a *shared* artificial world without normalization. None of the perspectives that can relate to and endorse the artificial social world would find themselves normalized away, for each would be related to the public artificial world in a way that makes sense *to* that perspective.[27]

In the Open Society, then, there is a kind of overlapping consensus on the public moral constitution amongst that society's 'reasonable' perspectives on ideal justice (that is, those perspectives that do not insist on adopting the 'optimizing stance' with respect to their ideal of justice, and instead conform to a norm of reciprocity in their interactions with other perspectives). There is no *single* conception of justice shared by all members of the Open Society, of course, but the elements that make up the public moral constitution—such as individual liberties, rules concerning property, markets, and so forth—provide a framework for social cooperation and interaction based upon a norm of reciprocity.

But what can advocates of *particular* conceptions of justice do politically within the Open Society? On this matter Gaus writes:

> [W]e must recognize that ideals of distributive justice are part of particular perspectives on justice, and in the Open Society no perspective has a special claim to have its ideal legally instituted. Questions of distribution, like so much, are *matters of democratic politics*. A democratic polity in the Open Society must beware of undermining the moral constitution that renders a shared public life among diverse perspectives possible, but it has many tasks that go *beyond* maintaining this general framework (Gaus 2016, p. 202, italics added).

This position seems entirely compatible with our account of the political liberal version of the well-ordered society and the place of justice as fairness within it. All reasonable political conceptions of justice must satisfy certain criteria, but which particular laws and policies are adopted within the PL WOS—including laws and policies concerning economic distribution —is a matter to be decided by citizens working within the procedures of the democratic constitutional structure of their society. Rawlsians, of course, will push for laws and policies that they take to be justified by the two principles of justice as fairness. But as Rawls himself emphasizes, because there are no 'philosophical experts' to decide these political matters for all in a democratic society ('Heaven forbid!'), they may do so only as 'citizens among others.'

III. THE CHOICE? WHAT CHOICE?

Having established that even within Gaus's Open Society Rawlsian citizens should be free to push democratically for the laws and policies that they think will move their society closer to their ideal of justice, it may remain the case that they still face The Choice. In this final section we try to cast doubt on Gaus's claim that The Choice poses a real problem for Rawlsian citizens.

Recall that Gaus maintains that ideal theories, if they are to constitute a distinctive approach to thinking about justice, *must* confront The Choice (Gaus 2016, pp. 82-84). Roughly, The Choice holds that adherents of a conception of ideal justice often must decide between: (a) pursuing 'nearby' gains in justice, with the risk that such local improvements may take their society further *away* from ideal justice (the full or most complete possible institutional realization of their favoured conception of justice); or (b) forego certain local gains in justice in order to pursue—at some greater distance, likely outside the immediate 'neighbourhood' of their social worlds, and thus more difficult for them to ascertain or see clearly—ideal justice. Drawing on his reading of Sen, Gaus denies that ideal theory is distinctive or helpful *if* it presupposes that the 'peak' of ideal justice is in our neighbourhood and/or that the 'terrain' of relevant social worlds is 'smooth.' In other words, if we face what Gaus refers to as a 'Mt. Fuji' terrain of social worlds—with ideal justice as the peak of the mountain, and all social worlds accessible to us, other than the ideal and our own, located either higher up or below us on the slope—then we do not need an ideal theory of justice at all. Instead, we can simply focus on climbing 'up,' as *any* movement upwards will get us closer to the top of the mountain. In contrast, if we confront a 'moderately rugged' region of social worlds (think of the Himalayas, with various peaks and valleys), then ideal theorizing can make sense, according Gaus, since simply climbing 'up' may lead us to the top of a peak *other* than the highest one, that is, a social world that is only somewhat just (according to our favoured conception of ideal justice), rather than fully just (or the closest to ideal justice possible for us).

(As an aside, we do not think that it is the case that ideal theorizing plays *no* helpful role even within a Mt. Fuji region

of possible social worlds. Consider two rival conceptions of justice: classical utilitarianism and justice as fairness. It may be that the terrain between our current social world and a social world that realizes (as fully as possible) classical utilitarianism is a smooth slope: any institutional changes we make to our society will either move us 'up' the slope towards the peak of classical utilitarian utopia, or move us 'down' away from it. Imagine that the same applies to the relation between our current social world and the realistic utopia of justice as fairness. Assuming that these are the two accounts of justice that we think are the best available options, we are confronted with two possible routes: climb up towards the classical utilitarian utopia or up towards the justice as fairness one. While we may go *some* distance up both slopes,[28] there inevitably will come a point, probably quite early in our journey, where climbing 'up' the classical utilitarian mountain involves climbing 'down' the justice as fairness mountain (and vice versa). While there are a number of reasons why we might opt in favour of, say, the conception of justice as fairness over classical utilitarianism,[29] and thus begin our climb without thinking about the peak, surely *one* consideration that might inform our choice are pictures (perhaps only sketches) of the two peaks in question, that is, views of what societies that fully realize the two accounts of justice in question look like. Consequently, ideal theorizing may be helpful even on Mt. Fuji.)

One noteworthy feature of Gaus's discussion of The Choice is how abstract it is. Very little is provided by Gaus in terms of concrete examples of foregoing local improvements in justice for the sake of pursuing the more distant goal of ideal justice. The historical examples of groups making The Choice mentioned by Gaus (as far as we could find) are limited mainly to twentieth century communist movements.[30] While we do not quarrel at all with Gaus's use of these examples, we very much doubt that they are applicable to Rawls's conception of justice as fairness (and reasonable political conceptions of justice more generally). If The Choice is a general problem, applicable to *all* ideal theory accounts of justice, then we think that it would have been helpful to see some cases where pursuing the kinds of political and economic institutions that Rawls takes to be recommended by justice as fairness—say, those that characterize what Rawls calls a "property-owning democracy"[31] —involve clearly *foregoing* significant 'local' improvements in justice in existing welfare-state capitalist societies.[32]

For the remainder of this section we will outline a way of thinking about local justice that naturally leads to what may be called 'full ideal theorizing.'[33] Our account constitutes an approach to promoting Rawlsian justice for which The Choice is not (we believe) a problem.

Recall that Rawls's criterion of reciprocity underpins his ideal of public reason. When citizens give public reason justifications for their political proposals, they necessarily are committed to the acceptability of those justifications for other citizens. Moreover, as we suggested earlier, there is an important connection between the acceptability of a justification for, say, a law, and compliance with that law. In finding a justification for a law acceptable citizens are willing, ceteris paribus, to *comply* freely with that law, that is, citizens acknowledge the normative authority of that law for their behaviour.[34] This is what it is for reasonable citizens to be motived by reciprocity in their public political relations with others. One reason for citizens to advance justifications for laws that satisfy the criterion of reciprocity is to bring about (or make possible)[35] compliance with those laws in a manner compatible with the deliberative agency of the citizens in question.

Now when citizens advance a political proposal for their society's basic structure—say, a proposal that they think will improve the justice of their society—part of the process of justifying that proposal is indicating what their basic structure would look like should that proposal be implemented. The hope is that other citizens will find the picture of the revised basic structure to be normatively attractive, and thereby come to support freely the proposal in question. Part of a public reason justification for a political proposal is the description of an alternative social world in which that proposal, through the (adequate) endorsement or acceptance of other reasonable citizens, is realized. Thus a public reason justification for a particular political proposal involves what we might call 'local ideal theorizing': consideration of an amended basic structure following reasonable citizens' acceptance of, and consequent compliance with, the political proposal in question.

Finding a justification for a political proposal convincing, and consequently endorsing that proposal as the most just or best proposal for one's society, differs from finding a proposal and its justification 'merely' acceptable (we made this point earlier with respect to reasonable political conceptions of justice within the PL WOS). What is essential for our account is that public reason justifications for political proposals are at least acceptable to all reasonable citizens. In finding a political proposal acceptable, citizens have adequate reason to *comply* with it.[36] It is part of the nature of a public reason justification that a successful justification —a justification that is (at least) acceptable to the reason-

able persons to whom it is addressed—will motivate compliance. This is because a reasonable citizen *cannot* both (a) find a public reason justification for a law acceptable, and (b) be *un*willing to comply freely (ceteris paribus) with that law. To affirm both (a) and (b) reveals a citizen to not be committed to reciprocity in her relations with others, and hence to be unreasonable. Thus public reason justifications, as they are addressed to reasonable citizens, involve at least local ideal theorizing.

While public reason justifications for political proposals cannot be severed from local ideal theorizing, what is the relation between local ideal theorizing and full ideal theorizing? How might citizens go from evaluating particular political proposals supported by public reason justifications to evaluating entire basic structures? Here is our suggestion. Some citizens endorse *multiple* political proposals. Such citizens judge their basic structure to be in need of wide-ranging reform. And each of these proposals aims at being acceptable to other citizens by means of public reason justifications; hence each of these proposals involve local ideal theorizing. Yet the proposals also should be *mutually realizable*, that is, citizens who endorse multiple political proposals for their basic structure should do what they can to ensure that their various proposals, should they be implemented, do not undermine or conflict with each other.[37] It makes little sense, at least normally, for citizens to aim at political proposals x, y, and z, if x undermines or conflicts with y, y undermines or conflicts with z, and x undermines or conflicts with z.[38] In other words, a desideratum, if not a criterion, of a set of political proposals is that they all be (at least) mutually realizable.[39]

At the limit, in offering a wide range of 'local' political proposals, all of which aim at acceptance by, and thereby compliance on the part of, other reasonable citizens, citizens may end up describing a realistically utopian version of their society. More precisely, through the process of determining how their various political proposals fit together and can be supported adequately by public reason justifications, some citizens *may* find it necessary or helpful to engage in something like Rawlsian ideal theorizing.[40] According to the account sketched in this section, though, such ideal theorizing can begin with a concern with *local* political reforms. By proceeding in this way, it is not clear why The Choice poses a distinct problem for such Rawlsian citizens.

The above discussion was rather abstract. We now will sketch how that account may work with reference to our current political and economic condition. Citizens naturally may reflect critically upon their circumstances, the basic structure in which they find themselves. In the case of the United States, they might conclude that its basic structure manifests a number of features incompatible with the ideal of democratic equality, understood as fair social cooperation among free and equal citizens. For instance, wealthy citizens have exercised, and increasingly exercise, highly disproportionate influence within the American political system.[41] Moreover, this influence has altered the basic structure of the United States in ways that have dramatically *increased* economic inequality over the past four decades.[42] Not only is this growing inequality economically damaging to society overall,[43] it has not improved the absolute incomes of the 'least advantaged' within the United States during this period, that is, there has been no noteworthy 'trickle-down' of economic benefit to the least advantaged.[44] Moreover, recent research on the intergenerational elasticity of citizens' incomes suggests "that the United States is very immobile," and thus falls far short of realizing anything like a principle of equality of opportunity (Mitnik and Grusky 2015, p. 4).[45] And despite important changes to the legal structure of society over the past five decades, profound race-based and gender-based inequalities in income and wealth, economic opportunities, and political influence continue to persist.

In response to these features of their society, citizens of the United States committed to the ideal of society as a fair system of social cooperation amongst free and equal citizens might propose changes to their basic structure. Such changes might include, inter alia, public financing of election campaigns, reforms to the provision of basic education and the distribution of higher education (so that the distribution of education counter-acts, rather than reinforces, existing class- and race-based inequalities),[46] a guarantee of employment for all citizens, ensuring universal heath care for citizens,[47] and limiting the total amount of wealth that citizens can inherit in order to counter-act the intergenerational concentration of wealth within a small portion of the population.[48] Measures that aim at promoting greater racial[49] and gender equality[50] within society also can be justified as necessary for securing the freedom and equality of all.

In advancing these kinds of political proposals, citizens aim at securing other reasonable citizens' acceptance of—and, if implemented democratically, willing compliance with—them by providing mutually acceptable justifications. But in formulating such proposals, and in trying to justify them, citizens might also try to determine how, and indeed

whether, the proposals and their justifications *fit together*. In other words, citizens might ask: can these proposals be justified and organized via coherent and compelling underlying principles which, in turn, express or are based upon an ideal of fair social cooperation among free and equal citizens? This process of justification, reflection, and revision naturally may lead some citizens to identify what they take to be the most reasonable political conception of justice, the conception that, overall, they think best expresses the ideals of free and equal citizenship, and of fair social cooperation. And part of this process can include, for those citizens, trying to think about what their society would look like should all of their main political proposals be realized.

Consequently, beginning with critical reflection on the *local* injustices that citizens identify with their basic structure, and coming up with political proposals for addressing those injustices in a manner compatible with the criterion of reciprocity, some citizens may find themselves eventually trying to ascertain what a realistically utopian version of *their* society would look like. Through this process some citizens might end up supporting the general features of a property-owning democracy,[51] the form of society that Rawls thinks is both feasible and can realize successfully the ideal of free and equal citizenship. But citizens need not stop there. Their commitment to property-owning democracy is not unconditional, but rather is a revisable one. It may turn out that some features of 'welfare-state capitalism' also can be used in order to reform the basic structure so that it more adequately satisfies the principles of justice as fairness.[52] In any case, the processes of addressing *both* questions of local justice and overall (full) justice need not involve fixating on one set of questions at the expense of the other; instead, citizens' reflections on both sets of questions can (and we think should) inform each other.

One way to understand Rawlsian ideal theorizing, then, is to see it as *beginning* with a concern with local injustices. According to this account, citizens can move 'up' from a concern with local injustices to a realistically utopian conception of their society—and this process, of course, can involve moving back and forth (or 'up' and 'down') between local improvements in justice and reflections on overall justice. Drawing on this account, and focusing on the particular kinds of political proposals that this process might generate (e.g., commitments to the public financing of election campaigns, universal health care, a social minimum for all citizens, and so forth), it is not at all obvious to us that The Choice will confront the reasonable Rawlsian citizens we are envisioning in any stark way.

Certainly there may be some cases in which difficult choices have to be made by political actors, and the long-term consequences of any particular policy or legislative decision can never be fully foreseen. But these difficulties do not seem peculiar to the kind of ideal theorizing that we have described here. All of the proposals sketched in the above paragraphs, for instance, are not only 'local' improvements in justice (at least from the perspective of justice as fairness), but *also* seem to move society closer to the institutions characteristic of a property-owning democracy (or at least do not appear to move society further *away* from property-owning democracy). Perhaps we are wrong about this. But appreciating the gravity of The Choice—or whether it even applies to many or most significant political decisions—requires more detailed consideration of what promoting a democratic and liberal conception of justice, like justice as fairness, involves in *practice*. Based on what Gaus says in *Tyranny*, we cannot see why the kinds of political proposals that would improve the justice of existing welfare-state capitalist societies (like the United States) plausibly would involve moving those societies *away* from the Rawlsian ideal of property-owning democracy.

IV. CONCLUSION

To conclude, we remain unconvinced that one main target of Gaus's critique—Rawlsian political liberals—are operating under a *tyranny* of ideal theory. While it is fair to say that justice as fairness requires revision in light of the insights and commitments of political liberalism, it is hardly a jumbled mess ('in disarray'). Moreover, Gaus's insistence that Rawls or Rawlsians owe an account of *the* liberal theory of justice is simply strange in light of his recognition that a commitment to (reasonable) pluralism entails a rejection of any such single, unified account as *the* account of justice.[53]

Here we have formulated a version of the idea of the well-ordered society—the PL WOS—that is consistent with political liberalism's commitments and aims. The PL WOS avoids Gaus's critique of Rawls's original formulation of the idea. We also have argued that, as it turns out, the Open Society that Gaus defends is not very dissimilar from the PL WOS. And finally, we proposed that there is a way to understand the role of public reasoning and ideal theorizing within political liberalism such that political liberals do not face (a serious version of) The Choice. Rather than being subjugated under the tyranny of ideal theory, then, perhaps Rawlsians are operating under the *democracy* of ideal theory.

NOTES

1 Gaus's use of the term 'Open Society' is an homage to Popper (1962), a book Gaus thinks has been undeservedly neglected and underappreciated by political philosophers.

2 Contrary to some recent work that argues that ideal theorizing is unnecessary or even counterproductive with respect to the aim of improving the overall justice of society (e.g. Sen 2009), Gaus proposes that ideal theorizing—*if* its aims are properly reconceptualised—in fact can help improve society's overall understanding of the requirements of justice. Ideal theorizing may be tyrannical in its *current* form, as practiced by many contemporary political philosophers, but it is not *necessarily* so, and in fact can play a constructive social role.

3 For Rawls's explanation of "ideal" and "non-ideal" theory, "partial compliance" and "strict compliance" theory, and the ideas of a "well-ordered society" and a "realistic utopia," see Rawls 1999, pp. 4-5, 7-8, 215-16, 308-9; Rawls 2001, pp. 4-5, 13, 65-66. The main assumptions of Rawlsian ideal theory are "strict compliance" and "favourable circumstances" (Rawls 1999, p. 216; 2001, p. 101).

4 Rawls (2005) initially articulates the idea of 'a comprehensive doctrine' by contrasting political conceptions of justice with moral doctrines (p. 13). He understands moral doctrines to be "comprehensive views" insofar as they have wide scope, covering a wide range of subjects. As such, a moral conception is comprehensive "when it includes conceptions of what is of value in human life, and ideals of personal character, as well as ideals of friendship and of familial and associational relationships, and much else that is to inform out conduct, and in the limit to our live as a whole" (p. 13). In contrast, "a political conception tries to elaborate a reasonable conception for the basic structure alone and involves, so far as possible, no wider commitment to any other doctrine" (p. 13). (On the relation between 'comprehensive doctrines' and 'conceptions of the good,' see Rawls 2001, p.19.)

5 The reason why we refer to our position as 'Rawlsian' rather than 'Rawls's' should become clear in Sec II.

6 In Rawls's final formulation of justice as fairness (2001, pp. 42-43), roughly, the first principle specifies a set of 'basic liberties' that are to be secured equally for all citizens within the constitutional structure of society (these liberties include freedom of thought, liberty of conscience, freedom of association, the political liberties (including their 'fair value'), freedom of political speech, and the like). The second principle requires that any economic inequality in society must (a) benefit the 'least advantaged' citizens over time more than any other system of economic distribution, and (b) not undermine or violate the 'fair' equality of opportunity of all citizens to compete for positions of authority and responsibility. The first principle, moreover, enjoys 'lexical' priority over the second.

7 The 'basic structure,' roughly, consists of the main institutions of society understood as an overall system of social cooperation. These institutions apply to all citizens within society, unlike 'voluntary associations,' such as religious organizations, firms, unions, clubs, and universities. (See Rawls 200), Sec 4; Rawls 2005, Lecture VII.)

8 For an insightful discussion of this development, see Weithman 2010.

9 Rawls emphasizes this repeatedly. See, e.g., Rawls 2005, pp. 68, 136, 216.

10 Rawls 2005, pp. 446-47, xliv-xlv. The criterion of reciprocity also grounds the "liberal principle of legitimacy" (Rawls 2005, pp. xliv, 137, 446-47). In fact, the criterion of reciprocity expresses the "intrinsic (moral) political ideal" of political liberalism (2005, p. xlv).

11 While there is "a family of reasonable political conceptions" of justice, "[t]he limiting feature of these forms is the criterion of reciprocity" (Rawls 2005, p. 450).

12 The basic structure, recall, is made up of society's main political and economic institutions, understood as an overall system of cooperation (see note 7).

13 For Rawls's final account of how this might be achieved, see his "Reply to Habermas" in Rawls (2005). (The achievement of such integration may involve *revisions* to citizens' comprehensive doctrines.)

14 See Gaus (2010, pp. 233-75; 2011; Gaus and Vallier 2009, pp. 51-76).

15 It should be noted that some Rawlsians hold that the idea of public reason should apply to *all* political decisions (see Quong 2011, ch.9). The question of the appropriate scope of public reason is not central to the differences between Rawls's and Gaus's views that concern us here.

16 In his "Reply to Habermas" (Rawls 2005), Rawls refers to this as the 'pro tanto' justification of a reasonable po-

litical conception of justice. Weithman (2011) contends that by not employing the idea of a freestanding political conception of justice, the convergence view of public justification cannot realize citizens' political autonomy; instead, in a society governed by the convergence account of public justification, citizens are politically heteronomous.

17 Throughout the book Gaus makes use of the idea of 'social worlds,' which he takes from Rawls. See, e.g., Rawls (2001, p. 128; 2005, p. 53).

18 Gaus takes this metaphor from Simmons 2010.

19 Neither sex nor race were originally included among the kinds of information deliberators in the original position were deprived of behind the veil of ignorance, and representative deliberators were said to be 'heads of households.' English (1977) was the first to criticize the heads of household assumption. Okin (1989, especially Ch. 5) famously criticized Rawls for failing to include 'sex' among the characteristics of which representatives in the original position lack knowledge. With publication of *Political Liberalism* Rawls explicitly enumerates sex and race, among other traits, as characteristics about which representatives in the original position lack knowledge (see Rawls 2005, p. 25; 2001, p. 15).

20 Thus strict compliance is *assumed* only within the first stage of the original position, specifically, the stage at which the parties initially select which conception of justice should govern the basic structure of the society in which the citizens whom they represent will live out their lives. The second stage of the original position involves determining whether a fully just well-ordered society—a society with a basic structure that is organized in accordance with the conception of justice selected at the first stage —would be stable over time for the right reasons, namely, through the free compliance of its reasonable citizens (see Rawls 2001, Sec 54-55). Here compliance is *not* assumed but must be demonstrated to be feasible, given the kinds of psychologies and interests that citizens can be expected to have in the society in question.

21 On the ambiguity of the nature of Rawls's concern with stability in *Theory*, see Gaus 2014.

22 Rawls claims that any conception of justice that fails to include these three features cannot satisfy the criterion of reciprocity within a pluralist society. We think that this claim is correct, though we cannot defend it here.

23 There may be some citizens who find two or more conceptions equally plausible, or who find all reasonable conceptions 'acceptable' but do not endorse any particular one (we explain what it is for a citizen to 'find acceptable' a conception in this paragraph). Such complications do not affect our discussion.

24 Because of their commitment to reciprocity, then, reasonable citizens do not adopt the 'optimizing stance' with respect to their favoured conceptions of justice (Gaus 2016, pp. 215-18).

25 It may be that different reasonable political conceptions of justice 'overlap' in justifying the same kinds of institutions within the basic structure.

26 Gaus is referring to Rawls 1958.

27 Our emphasis on 'endorsable,' 'endorse,' and 'shared.'

28 Indeed, an important aspect of Sen's work is that citizens can improve the overall justice of their society even if they do *not* share the same conception of justice. But obviously this will not always be the case.

29 We assume that we need not rehearse Rawls's arguments against classical utilitarianism here.

30 E.g., see Gaus 2016, pp. 88, 143. (And even these examples are mentioned only in passing.)

31 Rawls emphasizes the distinction between a property-owning democracy and a welfare state in the following way: "One major difference is that the background institution of a property owning democracy, with its system of (workably) competitive markets, try to disperse the ownership of wealth and capital, and thus to prevent a small part of society from controlling the economy and indirectly political life itself. Property owning democracy avoids this, not by redistributing income to those with less at the end of each period, so to speak, but rather by ensuring the widespread ownership of productive assets and human capital (educated abilities and trained skills) at the beginning of each period, all this against a background of the equal basic liberties and fair equality of opportunity" (Rawls 1999, pp. xiv-xv.)

32 For Rawls's discussion of five different 'ideal types' of socio-economic systems ('laissez-faire capitalism,' 'state socialism,' 'welfare-state capitalism,' 'liberal-democratic socialism,' and 'property-owning democracy'), see Rawls 2001, Sec 41-42, 49. Further discussion of the idea of a property-owning democracy can be found in Krouse and McPherson (1988) and O'Neill and Williamson (2012). (Most contemporary liberal societies, such as Canada, Denmark, Japan, and the United States, are welfare-state capitalist societies in Rawls's sense.)

33 This discussion draws upon Sec 2-3 of Neufeld (2017).

34 This claim concerns the way in which public reason justifications are meant to function in citizens' deliberations about fundamental political issues. We do not mean to presuppose anything controversial about the nature of 'reasons' per se (say, some form of 'reasons' or 'judgement' 'internalism').

35 Of course, compliance may often simply be the result of habit or deference to authority. What is important is that such compliance *could* be justified or, if necessary, be *brought about* via the rational agency of citizens.

36 It may be that finding acceptable or even endorsing a public reason justification for a law or institution is not always sufficient to bring about adequate compliance on the part of citizens with that law or institution. For instance, citizens may be subject to foreseeable akrasia even with respect to their compliance with laws and institutions that they support. In such cases, though, insofar as citizens accept the justification(s) for the law or institution in question, and thus agree that their compliance is warranted, they will accept (via instrumental reasoning) the use of those institutional mechanisms necessary to bring about and ensure their own compliance over time. (If the costs of such institutional mechanisms are quite high, though, reasonable citizens may reconsider their acceptance or endorsement of the laws or institutions in question.)

37 Citizens may find acceptable incompatible political proposals as alternatives, but in doing so they of course acknowledge that only one of these alternatives can be realized within their basic structure.

38 It would be rational, however, for citizens to promote political proposals *x*, *y*, and *z*, even if those proposals conflict with or undermine each other, *if* the citizens in question believe that (a) only one (at most) of those proposals has a chance of being implemented, (b) the implementation of any one of those proposals would improve the overall justice of their society, and (c) they do not know which one (at most) of *x*, *y*, or *z* will (possibly) be implemented. Such unusual cases can be put aside for our purposes here.

39 It also may be desirable that the proposals support or reinforce each other. Here we focus on the weaker claim that they should at least be mutually realizable.

40 This is not to say that in endeavouring to ensure the coherence of their various local political proposals citizens *must* turn to Rawlsian ideal theorizing (with its focus on the basic structure as an overall system of social cooperation, and so forth), only that it is *one* plausible strategy for doing so. (Thanks to an anonymous commenter for pressing us to clarify this point.)

41 See Gilens and Page (2014).

42 See Hacker and Pierson (2010). On the overall increase in income inequality within the United States in recent decades, see: Congressional Budget Office 2009, 2011.

43 On the economic harms of high levels of inequality in income and wealth, see: Stiglitz 2012 and Galbraith 2014.

44 The main reason for this, according to Kenworthy (2010), has been government policy decisions.

45 See also Isaacs, Sawhill, and Haskins (2008).

46 Among such educational reforms would be those that promote racial integration (see Anderson 2010).

47 The proposals concerning campaign financing, employment, and heath care are mentioned in Rawls (2005, pp. xlvi-xlvii).

48 On the importance of minimizing the intergenerational accumulation of wealth within a small number of citizens, and the consequent role of taxing bequeathments and gifts, see Rawls (1999, pp. 245-246; 2001, pp. 160-161. Piketty (2014) recently has documented the long-term tendency of capitalist societies toward what he terms 'patrimonial capitalism.' A patrimonial capitalist society, roughly, is one in which the members of that society's economic elite enjoy their privileged position primarily as a consequence of inheritance, not innovation or entrepreneurship. Simplifying greatly, the reason for this tendency is that returns to capital ('r') generally grow at a higher rate than the overall economy ('g'). Consequently, the already wealthy within society tend to become wealthier at a much faster rate than anyone else, and, moreover, pass this advantage on to their descendants. This economic elite becomes largely a class of rentiers. The members of this class also are able to employ their wealth to influence the political decision-making processes of their society, thereby undermining the democratic equality of citizens. Piketty's research appears to support Rawls's more speculative worries about the long-term tendency of capitalist societies toward growing inequality, decreasing political freedom for most citizens, and hence injustice.

49 On extending Rawls's account of justice to address racial inequality, see Shelby (2004; 2013). (The account of ideal theory that we advance here is, we believe, broadly compatible with Shelby's position.)

50 For public reason justifications for laws and policies that promote gender equality, see: Baehr (2008),

Brettschneider (2007), Hartley and Watson (2010); Lloyd (1998); and Neufeld (2009).

51 See n.31.

52 Krouse and McPherson (1988) suggest this with respect to the role of a progressive income tax in promoting justice as fairness. O'Neill (2012) recommends that we regard the institutions of both property-owning democracy and welfare-state capitalism as comprising a kind of 'toolkit' that egalitarians (citizens committed to justice as fairness) can draw upon when trying to improve the justice of their society.

53 Gaus (2016, pp, 153-154, n. 8) writes, anticipating one kind of political liberal reply, "Some might argue that political liberalism is concerned with legitimacy, not justice. Even if so, this would not show that a coherent *theory of justice* remains. […] 'It's only about legitimacy' is not a magic phrase that can make these issue disappear. What *is* the liberal theory of justice?"

REFERENCES

Anderson, E. (2010). *The Imperative of Integration.* Princeton: Princeton University Press).

Baehr, A. (2008). Perfectionism, Feminism, and Public Reason. *Law and Philosophy* 27:193-222.

Brettschneider, C. (2007). The Politics of the Personal: A Liberal Approach. *American Political Science Review* 101:19-31.

Congressional Budget Office (2009). Data on the Distribution of Federal Taxes and Household Income.

Congressional Budget Office (2011). Trends in the Distribution of Household Income between 1979 and 2007.

English, J. (1977). Justice Between Generations. *Philosophical Studies* 31:91-404.

Galbraith, J. (2014). *The End of Normal: The Great Crisis and the Future of Growth.* New York: Simon & Schuster.

Gaus, G. (2010). Coercion, Ownership, and the Redistributive State: Justificatory Liberalism's Classical Tilt. *Social Philosophy & Policy* 27: 233-75.

— (2011*). The Order of Public Reason.* Cambridge: Cambridge University Press.

Gaus, G. (2014). The Turn to a Political Liberalism. In: D. Reidy and J. Mandle (eds.) *The Blackwell Companion to Rawls.* New York: Wiley-Blackwell, pp. 235-50.

— (2016). *The Tyranny of the Ideal: Justice in a Diverse Society.* Princeton: Princeton University Press.

Gaus, G. and Vallier, K. (2009). The Roles of Religious Conviction in a Publicly Justified Polity: The Implications of Convergence, Asymmetry, and Political Institutions. *Philosophy & Social Criticism* 35: 51-76.

Gilens, M. and Page, B. (2014). Testing Theories of American Politics: Elites, Interest Groups, and Average Citizens. *Perspectives on Politics* 12: 564-81.

Hacker, J. and Pierson, P. (2010). *Winner-Take-All Politics*. New York: Simon and Schuster.

Hartley, C. and Watson, L. (2010). Is a Feminist Political Liberalism Possible? *Journal of Ethics and Social Philosophy* 5:37-54.

Isaacs, J, Sawhill, I, and Haskins, R. (2008). Getting Ahead or Losing Ground: Economic Mobility in America (The Economic Mobility Project, the Brookings Institution and the Pew Charitable Trusts).

Kenworthy, L. (2010). Rising Inequality, Public Policy, and America's Poor. *Challenge* 53:93-109.

Krouse, R. and McPherson, M. (1988). Capitalism, 'Property-Owning Democracy,' and the Welfare State. In: A. Gutmann (ed.) *Democracy and the Welfare State*. Princeton: Princeton University Press, pp.79-105.

Lloyd, S. (1998). Toward a Liberal Theory of Sexual Equality. *Journal of Contemporary Legal Issues* 9:203-24.

Mitnik, P. A. and Grusky, D. B. (2015). Economic Mobility in the United States. The Pew Charitable Trusts and the Russel Sage Foundation.

Neufeld, B. (2009). Coercion, the Basic Structure, and the Family. *Journal of Social Philosophy* 40: 37-54.

— (2017). Why Public Reasoning Involves Ideal Theorizing. In: K. Vallier and M. Weber (eds.) *Political Utopias: Contemporary Debates*. Oxford: Oxford University Press, pp. 73-93.

Okin, S. M.(1989). *Justice Gender and the Family*. New York: Basic Books.

O'Neill, M. and Williamson T. (eds). (2012). *Property-Owning Democracy: Rawls and Beyond*. Oxford: Wiley-Blackwell.

Piketty, T. (2014) *Capital in the Twenty-First Century*. Cambridge MA: Harvard University Press.

Popper, K. (1962). *The Open Society and Its Enemies*. London: Routledge and Kegan Paul.

Quong, J. (2011). *Liberalism without Perfection.* Oxford: Oxford University Press.

Rawls, J. (1958). Justice as Fairness. *The Philosophical Review* 67: 164-94.

Rawls, J. (1999). *A Theory of Justice*. Revised Edition of 1971. Cambridge, MA: Harvard University Press.

— (2001). *Justice as Fairness: A Restatement*. Cambridge, MA: Harvard University Press.

— (2005). *Political Liberalism*. Expanded Edition, Original edition 1993. New York: Columbia University Press.

Sen, A. (2009). *The Idea of Justice*. Cambridge MA: Harvard University Press.

Shelby, T. (2004). Race and Social Justice: Rawlsian Considerations. *Fordham Law Review* 72:1697-1714.

— (2013). Racial Realities and Corrective Justice: A Reply to Charles Mills. *Critical Philosophy of Race* 1:146-62.

Simmons, A. J. (2010). Ideal and Nonideal Theory. *Philosophy & Public Affairs* 38:5-36.

Stiglitz, J. (2012). *The Price of Inequality: How Today's Divided Society Endangers Our Future*. New York: W.W. Norton.

Weithman, P. (2010). Why Political Liberalism? On John Rawls's Political Turn. New York: Oxford University Press.

— (2011). Convergence and Political Autonomy. *Public Affairs Quarterly* 25/5: 327-48.