How Can We do Political Philosophy?

FRED D'AGOSTINO

President of the Academic Board
Office of the President of the Academic Board
The University of Queeensland
Brisbane St Lucia, QLD 4072
Australia

Email: f.dagostino@uq.edu.au
Web: http://researchers.uq.edu.au/researcher/1345

The Tyranny of the Ideal, by Gerald Gaus (hereafter *TI*), is an examination, in two parts, of the prospects for a type of political philosophy that provides direction for reform while remaining in contact with the contingencies, diversities, and infirmities of the human condition.[1]

In the first, longer part, which is critical in intention, Gaus considers the viability of an approach, labelled *ideal theory,* that has been influential for millennia and concludes, to put it crudely, that this approach is misbegotten, in particular, because of the complexity of the relations among the congeries of factors that are relevant to defining a political ideal, including, in particular, the diversity of individuals' perspectives on their good and the good for their society. My comments on this part of the book will largely take the form of a reconstruction, via a different mode of presentation, of some of Gaus's main points. The key idea here will be that the problem of political philosophy is cast within a framework of *layered complexity.*

In the second part, which is constructive in intention, Gaus, first, offers an alternative approach, couched in terms of Karl Popper's idea of "the open society" as a privileged configuration of social arrangements that is designedly friendly to precisely the forms of diversity that, on Gaus's account, undermine the ideal theory approach, and, second, essays the prospects that such a configuration might be a point of convergence for individuals with (as it turns out, not quite) the full range of perspectives. One of my critical points will be about the prospects for "the open society" in current socio-political circumstances, characterized, as they seem increasingly to be, by various forms of polarization, fundamentalism, denialism, extremism, irrationalism, and rejectionism.[2]

These matters have an importance beyond the immediate argumentative dialectic, or so I believe. In particular,

Gaus's approaches, both critical and constructive, point towards a new way of doing political philosophy, one that has been emerging in his own work and, increasingly, in the work of others, such as Ryan Muldoon's (2016) recent *Social Contract Theory for a Diverse World: Beyond Tolerance* and Julian Müller's forthcoming *Capitalizing on Political Disagreement: the Case for Polycentric Democracy.* So, as a way of framing *TI,* I begin by saying something about this emergent new program in political philosophy.

THE NEW PROGRAM IN POLITICAL PHILOSOPHY

To speak of a "new" program in political philosophy is to imply the existence of an existing or established or simply "old" program and that, of course, is the Rawlsian one, initiated, for all intents and purposes, by the 1971 publication of the book *A Theory of Justice,* which, along with its various successors, has been a point of focus for Anglosphere or more broadly "analytic" political philosophy ever since.[3] Indeed, both Gaus and I have contributed to that very program of research. But less reluctantly than others, or maybe it's just more recklessly, we have concluded that the Rawlsian project has failed, to adopt the terminology of footnote 1, to effectively manage "the essential tension" between the diversity of individuals' interests, values, "comprehensive conceptions of the good" and so on (all acknowledged by Rawls) and the necessity, as Rawls saw it at least (though we do not), that a public conception of justice provide a basis, grounded in a deep moral consensus, for the coordination of social action for mutual benefit. As Gaus puts it (*TI,* 153-4):

29

Rawls insisted that a theory of justice was characterized by choice from a certain normalized perspective, but his later view allows multiple partially normalized perspectives that yield different conceptions of justice. However, if one acknowledges that there are other reasonable normalizations that yield inconsistent conceptions, in what sense can one plausibly claim that one has identified *the* principles of justice for the definitive ordering of social claims in a well-ordered society…?

Indeed, what seems to have happened is that Rawls abandoned, during the course of his long career, both the goal he set for political theory and the fundamental modelling device that he adopted as a basis for pursuing that goal. In particular, Rawls in effect abandoned the idea that political theory ought to and could successfully aim at the identification of a public conception of justice fit to order competing social claims. (This is the idea of the "well-ordered society".) And he also abandoned what had been a key, and highly arresting, heuristic for pursuing that goal, namely, the normalization of stakeholders, or, in particular, the abstraction, during theorizing, from their various differences, empirically—in their values priorities, in their social roles, in their "identities" (e.g. race, gender) etc. (This is the device of "original position" argumentation, where the original position is a unique privileged instance of a collection of initial situations of choice.) What we see here, in the reversal or abandonment of key features of Rawls's theoretical approach, is what one of my teachers, Imre Lakatos (1970), called a "degenerating research programme". So, as Gaus now puts it (*TI,* 153), "[o]ne 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."

Of course, Thomas Kuhn (1970, p. 151) already argued, indeed around the same time as *Theory* was published, that a well-entrenched approach to knowledge-making or theory-building would persist, even in the face of internal and external difficulties, so long as there were no reasonably well-delineated alternative to or replacement for it. Perhaps that explains the persistence of work on the Rawlsian project. But such an alternative is, arguably, now coming to hand, in the form of a new approach that Gaus has played a conspicuous role in developing.

One feature of *the new program* is its treatment of the empirical diversity of ethico-political stakeholders, which it simply accepts as a given and works with without any very elaborate abstraction or idealization in the interests of homogenization. Indeed, Gaus and others (including Muldoon, Müller, and me) all see precisely this diversity as a resource for the organization of social life, providing, as it does, the basis for a division of labor in the economy and for the better understanding of partially shared values and cognitive perspectives (through their encounters with different variants on common themes). So, whereas Rawls starts with *stakeholders are different so let's normalize their differences in the interests of constructing a consensus on the principles of social coordination*, the new program starts with *stakeholders are different so finding principles of social coordination will depend on finding points of convergence from an accepted diversity of unnormalized perspectives.*

Indeed, this way of putting the matter suggests a second conspicuous feature of the new program, namely its preference for convergence rather than consensus approaches to social coordination. This represents a real sticking-point for many participants in the Rawlsian project, who worry that the lack of a shared substantive basis for coordination is ipso facto lacking in normative authority or stability, and, in fairness, this issue does constitute a challenge for the new program that ought to be, indeed already to some extent is, on its to-do list (or, as Lakatos would put it, is part of its "positive heuristic"). See, for example, Kevin Vallier's (2014) recent book *Liberal Politics and Public Faith.* Because of the emphasis on convergence rather than consensus, we can expect, and indeed find, that sponsors of the new program put emphasis on argumentative or more broadly justificatory mechanisms that are convergence-compatible, such as

- bargaining, which, as Muldoon, following Hayek, points out, "has the advantage of requiring very little agreement between the parties involved" (Op. cit., p. 69);
- path-dependent social-evolutionary processes, as sketched by Gaus, for example, which can discover and stabilize a coordinating equilibrium even in the presence of considerable divergence of underlying motivations among participants;[4]
- reconciliation by separation or the recognition of separate spheres, as for example in systems of so-called jurisdictional rights, as analysed by Gaus (*TI*, IV.2.4);
- the appeal to the gains to trade, broadly understood, that are available via a division of labor built on diverse perspectives, as for instance, in an economistic sense in Muldoon's analysis (Op. cit., sec. 5.3) and more broadly by Gaus, especially in the form of his "Fundamental Diversity Insight" (*TI*, 133), and including in particular,

those gains resulting from Millian "experiments in living",[5] a concept which has become much more common in the political-philosophical literature of recent times.

A third important feature of the new program is its general understanding of the political-philosophical enterprise. In particular, the new program adopts what I call an "engineering" approach to political philosophy, seeing the goal of the enterprise as the design of devices or technologies (see *TI*, 183) that are fit for the purposes of ethico-political stakeholders in their pursuit of workable social arrangements. There is convergence from seemingly unrelated spheres of philosophical enquiry on this understanding, including, for instance, and indeed as an especially vivid example, the recent work of Elijah Millgram, who characterizes his own discipline, metaphysics, as "a design science (like architecture, or computer science, or mechanical engineering)" and who explains, at greater length, why we might better understand certain philosophical enterprises "as intellectual ergonomics". Millgram (2015, p. 15) says:

Throughout its history… metaphysics has been the answer to the question: how do we have to understand the world for reasoning about it to be possible? There is a *practical* spin to put on the question: how can we make reasoning and inference tractable and effective? One approach to the question so understood will focus on the design—and on the redesign—of intellectual devices that make it feasible to think about the world through which we must navigate ourselves. If that is what we are interested in, when we are doing metaphysics, then we should be attending not in the first place to what expressions *mean,* but to what the devices *do.*

(Others, including Muldoon and Gaus's student Chad van Schoelandt (2015) have all explicitly endorsed this approach.)

Perhaps most notably, though this point remains to be fully explicated, the new program more or less abandons the project of providing an end-state description of a justified social order, preferring, instead, and entirely in keeping with the engineering conception, to see its job as that of identifying a *method* of (i.e. tools for) thinking about the problems of political and social order,[6] rather than of deriving or otherwise justifying descriptive statements (of normative intent) about that order. While it cannot be precluded (though it also can't be guaranteed) that the new

approach, carefully applied, will provide some guidance to stakeholders, what it will not provide is a substantive account of what their social arrangements ought to be. That account will emerge from *their* activities, perhaps using the tools identified by the philosopher. Political philosophy, on this account, lets go of what the sociologist Zygmunt Bauman (1987) calls the "legislative" aspiration and cedes the legitimation of their social arrangements to the stakeholders who make those arrangements. And, although this is radically overdetermined, this is so, at least in part, because, unlike the situation in other forms of knowledge-making such as science, there is, despite a charming faith to the contrary widespread in the philosophical community, no external source or standard against which to measure the success of a process of reasoning or evolution in delivering a legitimate system of social arrangements. All that matters, really, is how that arrangement came into being… and that is what a theory of method in political philosophy can give us.[7]

So much for the circumstances in which *TI* might be understood. How about its main arguments? Let me begin with the critical or negative part of Gaus's development of ideas.

THE ENIGMA OF IDEAL THEORY

I have already mentioned the idea that political philosophy might aspire, indeed has typically been seen to aspire, to the derivation or justification of a description of an ideal social state. For political philosophy to be practical, even in a restricted sense that does not entail an engineering approach, the notion of an ideal social state has to include, at some level of granularity, a description of the specific institutional and cultural forms that sustain the realisation of, if you will, the social ideal. And the social ideal will be given, in turn, by some account of the ways in which, at least in typical situations and perhaps at a high level of abstraction, the various social goods are balanced optimally against one another (so that, crudely, there is no other balancing which would deliver net-gains in some appropriate way).

This is *ideal theory,* then, and, in the first 150 or so pages of *TI,* Gaus undertakes to expose its pretensions. I should say, at the outset, that I found Gaus's critical points entirely persuasive and have, as far as I can see, no corrections or amendments to suggest. What I would like to do, however, is present an account from a different perspective that will track Gaus's at least in its more important aspects. Perhaps this account will open up some points for my readers, as it

did for me. Let me begin by describing a model of *layered complexity* in relation to social order. (I note that this model in my view survives, in its main features, any critique of the pretensions of ideal theory and hence remains available for appropriation by advocates of the new program.)

At the base, we have a set **A** of arrays of social institutions, $\{A_1, A_2,\ldots, A_n\}$, where each array A_i consists of a system of specific institutional and cultural elements $<I^i_1, I^i_2,\ldots I^i_n>$, where the relations among the elements within a given array vary from loosely to tightly coupled.[8] Relations between different arrays are relations of similarity and difference, so that, for example, one array might differ in one particular institutional or cultural element from another and hence could be considered quite similar and indeed perhaps as quite accessible if we were thinking of social transformations. (It will be part of Gaus's argumentative strategy to assume, *arguendo* and as ideal theorists probably must assume, that we can define a similarity metric Σ over the set **A**. It will, of course, be a partial ordering rather than a complete one.)

One of the things we know about such arrays, when they are even moderately "realistic" in relation to our own social circumstances, is that it can be very difficult to understand all the relations between or among elements. Given coupling, how does changing one element affect others in the array? How, because of the complexity of the interrelations among elements, does changing one element affect the configuration of the array as a whole? We know the difficulty of these questions, purely empirical though they be, through the concept of "unintended consequences", already familiar to Adam Smith, but put into play in our own time by the sociologist Robert Merton (1936). Because or insofar as the relations among elements of an array are obscure or complex, changing an element can have consequences that we cannot predict and hence will not anticipate.

I have, on other occasions (D'Agostino 2010, esp. Ch. 7), written about complexity and so won't repeat the exposition, but the key point is that any given element in an array might be coupled with others so that changing it changes them (or vice versa) and not always in smoothly linear ways. For example, changing some element might have different kinds of effects on different elements with which it is coupled. Perhaps increasing the stringency of some social rule reduces a certain form of associated activity while also increasing another, also associated activity. All this is purely empirical, how changes to one element affect others… it is a matter for observationally informed social science, difficult, probably, to model in a purely theoretical sense,

except, perhaps, by simulations (perhaps involving agent-based modelling).[9]

The base level or first layer consists, then, of a set of arrays, each of which is a complexly organized collection of social institutions and cultural elements.

At the "top" of the layered complexity model is an evaluative instrument which identifies a number of metrics, μ_1, $\mu_2,\ldots, \mu_n$, against which the various arrays can be assessed.[10] Again, these metrics will be, variously, tightly and loosely coupled with one another, so if we are at all "realistic" relative to what we know of our own systems of evaluation, it will be characteristic that the various metrics will interact complexly with one another, so that changes (in institutional arrangements, of course) which increase value against one of these measures may well decrease value against others of them. (A change to a legal rule may increase liberty but decrease equality because of the way the effects of that change propagate through an existing system of social institutions.)[11] We may also, of course, have what I will call a reconciliation $\prod$ of these various metrics, trading them off against one another in some appropriate way that yields an *overall* assessment of the given arrays. And, given the complex coupling of metrics with one another, we can expect overall assessments to exhibit the usual nonlinearities.

For example, starting with the array A_p, perhaps we believe, even rightly believe, that changing the element I^p_j will increase the value of the resulting array against the measure μ_i. Given the way the metrics are coupled, this may not result, however, in an overall increase in value. Perhaps increases against μ_i are coupled with decreases against μ_j such that, because of μ_i/μ_j trade-offs, the $\prod$ value of the new array is lower than of the starting-point array. That is one complexity. But there is a second and it's internal to the empirical layer in our scheme. Remember that changing I^p_j may change other elements in the given array, perhaps, in this case, I^p_k. But now, even without the complexities at the evaluative layer, evaluating the effects of the I^p_j change are more obscure because any evaluation has to identify the I^p_k changes as well and consider how those changes affect the μ_i value of the array, and, of course, it cannot be guaranteed, in the face of complexity, that there is overall increase in μ_i value, given that the positive I^p_j changes might well be outweighed by negative I^p_k changes.

Returning to the idea of an ideal theory, we can see now, I think, that it will, just as Gaus proposes using other terminology, have two key elements… though notice how these must be linked. One element is the description, in terms of institutional elements, of the ideal social state, defined, of

course, as that member of the array **A**, A* for which the Π value, π* is maximized. To make this theory practical, we rely on the similarity measure Σ, so that we can determine, for arbitrary A_i, how distant it is from A* or, more crudely, how many changes to institutional elements, are required to transform A_i into A*.

Such a theory would tell us, then, what the ideal social state is like and would, at least in principle, imply a pathway between our existing social situation and that ideal social state, for example, by a series of small changes to the institutional elements, one by one, until all the differences between A_i and A* has been eliminated.[12] (Crucially, because of the nonlinearities, there may be no clear and consistent pathway from A_i to A*; the first step in a given direction may be value-increasing and distance-decreasing, whereas the next step, in the same direction, might be, again, distance-decreasing but value-decreasing.)

In calling this section of my paper "The *Enigma* of Ideal Theory", I meant to refer, in fact, to the notorious German war-time Enigma coding machine, the one whose workings were, according to the mythology, unpacked by Alan Turing. And I chose this locution because, in fact, the relation between a change in some institutional element and the overall evaluation of the resulting array is as obscure, in most reasonably "realistic" cases, as the relation, in the Enigma machine, between the letters typed in and the letters that were then lit up on the board, having passed through various enciphering devices within the machine. And that, in essence, is Gaus's argument about ideal theory. Because of the two complexities identified above (among others), there is no realistic prospect that we will be able to use the similarity measure as a guide to realising the ideal social state… the array whose overall value is at a maximum, the array that best realises the social ideal. Using the idea of "high-dimensional landscapes" to refer to what I've called the complexity of couplings, Gaus puts it (*TI,* 68-9) like this:

> In terms of our ideal theory model, in a maximally high-dimensional landscape there is no systematic relation between the justice of social world *i* [my A_i] and the justice of the worlds that are adjacent to it. Note that in such a landscape there is no point in getting *close* to the ideal point, *u* [my A*], but not achieving it: its near neighbors may not be at all just.… The crux of maximally high-dimensional landscapes is that the justice of any one rule or institution is a function of all others, producing what Kauffman called "a com-

plexity catastrophe."… [O]ur concern here is a political theory that seeks to judge the justice of various social worlds, and recommends moves based on its evaluations of these worlds. In this context, the idea of a complexity catastrophe is entirely apropos, for the system will be too complex—really chaotic—for the theory to generate helpful judgments and recommendations.

While there are other, indeed fascinating, bells and whistles, this is the crux of Gaus's critical analysis of the situation of political philosophy. If political philosophy is defined by an ambition to develop an ideal theory that meets certain reasonable criteria, then it is bound to fail and, accordingly (*TI,* 245-6), there is need for "a break with contemporary social and political thought", because the idea on which it is based, that of ideal theory, "is ultimately a mirage, yet one that tyrannizes over our thinking and encourages us to turn our backs on pressing problems of justice in our own neighbourhood."

Of course, it is one thing to undermine the legitimacy of the project of ideal theory and quite another to reject the model of layered complexity that in fact enables that demonstration to be undertaken in a rigorous way. While it may not be possible, any longer, to take seriously the idea that we can use the similarity measure Σ to set an unwavering course towards the ideal social state, A*, that we can take with confidence,[13] it is still possible, indeed highly desirable, to use the apparatus of empirical and evaluative layers as an aid to theoretical enquiry about the project of political philosophy. And, in particular, one of the points that emerges quite clearly from the contemplation of this model is that it will be highly relevant to any respectable project of political philosophy to learn more about the variety of institutional arrays, the relations among institutional elements that are characteristic of the more common arrays, the issues around transformation of one array into another, and so on. All these are empirical questions, not to be answered simply by conceptual analysis, although they might well be modelled, for instance through simulation exercises. Using Lakatos's terminology, the injunction to study these arrays, for instance through institutionalist enquiries in political science,[14] is a reasonable element of the "positive heuristic" for the new program in political philosophy. Insofar as the PPE project—the project of recognizing the complementarities among the three key disciplines of politics, philosophy and economics—is now reasonably well consolidated,[15] it perhaps could be extended by the more systematic inclusion

of institutionalist analyses, coupled, perhaps, with studies in political sociology and indeed political social psychology.[16]

I said, earlier, that there was a positive, as well as the reported negative part to Gaus's analysis. Let me turn, briefly, to that account, framed, as it is, largely in terms of Popper's idea of "the open society", another one of Gaus's targets for intellectual rehabilitation.

THE OPEN SOCIETY AND ITS EXTREMITIES

If the critical argument of *TI* establishes that ideal theory is not viable as a paradigm for political-philosophical enquiry, then the positive argument is meant to provide an alternative approach, one that, as indicated earlier, starts from the inescapable and indeed potentially beneficial fact of diversity and seeks, via an engineering approach, to establish the possibility of, and conditions for, a convergence on nontrivial principles of institutional design. This takes the form of what Gaus, following Popper, calls "the open society", or, in particular, a social state that is characterized by certain key features that enables its stakeholders to, as Gaus puts it (*TI,* 176), "live under a shared moral framework that is diversity accommodative because it is accommodative to diversity per se".

The open society, on Gaus's account, is not primarily characterized by what we might call fine-grained, descriptively rich, action-guiding normative injunctions. Because there is no ideal social state (the critical part of the argument), there can be no legitimate array of injunctions (and other cultural elements) of this particular kind. (Recall that the so-called new program abjures such a "legislative approach"; the failure of ideal theory is a reason.) What there might be, however, is a collection of (meta-level) principles that will provide design criteria for any array of concrete institutions that can claim to exemplify the open society. The new program, then, consistently with its proceduralism, and with its orientation to methodology rather than theory, delineates a set of abstract principles that will have to be instantiated (though in any of a variety of ways) in order to be "diversity accommodative" in a society which cannot agree about the social ideal, but which still seeks to obtain the benefits of mutuality. This is, I believe, what Gaus calls (*TI,* 149) "a moral, liberal framework… which abjures the pursuit of the ideal while providing a framework for diverse individual perspectives on justice."

And the key meta-level design principles for such an open society include, according to Gaus, at least the following:

- That there be opportunities for an individual to interact (a) with others who share their particular interpretation of the social ideal within a "republican" sub-community (*TI,* 146-7) and (b) with those who have different interpretations of the social ideal, typically via their shared interest, despite these differences, in some common practical problem (*TI,* 185).[17]
- That the substantive normative injunctions, whatever they might turn out to be, be understood, by all relevant parties, via what Gaus calls (*TI,* 187) "the principle of natural liberty", or, in other words, the meta-level interpretive principle that "[w]hatever is not prohibited… is permitted", on the grounds, specifically (*TI,* 196) that "a natural liberty system encourages experimentation and discovery" and hence is diversity accommodative.
- That prohibitions rather than permissions be the default form for specific normative injunctions (*TI,* IV.2.3.3) on the grounds that, in conjunction with the principle of natural liberty as a "closure principle", this will, again, encourage discovery and innovation because (*TI,* 196) "[m]oral experimenters… need not first convince themselves that a new action falls under a previous permission".
- That the system of concrete normative injunctions makes appropriate use of so-called "jurisdictional rights", which (*TI,* 199, 200), for participants in an open society, "decouple the perspectives and so lessen the complexity of the system", thus "allowing high levels of change in some perspectives without affecting the shared public world".
- That market transactions and exchange more generally be a primary modality of social coordination, since, as Gaus puts it (*TI,* 203), they "provide bridges between different perspectives [because although] each sees the object in different ways… [they] typically share enough so that… they can agree on what is being traded, and that each is better off".
- That specifically legal rules be relied on (*TI,* 207) "to provide shared classifications of prohibited behaviour among those with, essentially, different perspectives", thus permitting coordination of expectations and behaviour among a diversity of individual stakeholders.

As you can see, these are, as foreshadowed, meta-level design principles, rather than specific and concrete social ideals or behavioural norms. If we undertake to deliver something positive by way of political philosophy, it will take this form. If we imagine ourselves considering how to create an open society, in the absence of, and indeed on ac-

count of the absence of an agreed social ideal (with its corresponding ideal social state), then we should, on Gaus's account, do so by seeking, in particular concrete circumstances, to identify or create an array of social institutions which meets these meta-level requirements. Indeed, reverting to our set of arrays **A,** we see, I think, that these design principles will impose a partition on that set between the arrays which do and those which do not honor them. While this does not tell us, concretely, what to do to improve our social situation, it does tell us how to begin… by moving, from wherever we are now, towards an array in which there are jurisdictional rights, exchange relations, and a principle of natural liberty that provides a closure rule over specific prohibitions. Of course, because of the layered complexities considered earlier, our pathway may not be a uniformly smooth one, and we will want to proceed, as Popper recommends under the heading of "piecemeal social engineering" (Popper 1962), by small and ideally reversible steps.

Of course, in speaking of these meta-level principles as *design* principles, and, indeed, in speaking, earlier, of the *engineering* approach, I seem to commit myself to a position which, actually, I do not accept and which Gaus doesn't either. So, in particular, and to avert to the not entirely co-incidental title of the journal in which this article appears, I seem to be presupposing that, in our political philosophizing, we are thinking about the characteristics of a made order, a taxis in Hayek's terminology, rather than a cosmos, or grown order.[18] I am not. Sometimes, of course, engineers look to the products of untheorized practice to identify principles that can be articulated, systematised, and then rigorously tested. These design principles might well be, on this account, the deliverances of a reflective examination, by political philosophy, of the various grown orders revealed by history and political sociology. To develop some more sophisticated account of the relation between so-called design principles and grown orders is, I believe, a project for the new program, and one worthy of endorsement by the positive heuristic of that program.

Let me conclude by way of Gaus's forthright approach to what he calls (*TI,* 208) "the critical question of justification", one which is especially acute because, absent the impossible dream of an ideal social state, there are a diversity of perspectives on the social and because justification consists, in the broad framework in which Gaus works, in finding some surface of convergence (on the open society) from all (or most) of these various perspectives.

In fact, as Gaus is quick to recognize (*TI,* 215), whether there *is* a surface of convergence depends more or less cru-

cially on stakeholders "not insist[ing] on… the 'optimizing stance'—that the only rule that is acceptable is one's top-ranked rule." It is on this point that I want to focus, on account, specifically, of the prospect that more people are more inclined now to insist of precisely that, though not, perhaps, in exactly those terms (which are, after all, terms of art).

In particular, I believe that it is an open, sociological question—a question about the middle level in our model of layered complexity, the layer of social agents—whether individuals who differ in their interpretations of the social ideal, who would, upon reflection, offer different, especially opposed, accounts of the ideal social state are now able, in sufficient numbers, to recognize each other, nevertheless, as potential partners in a social contract to form an open society. This, I think, is the lesson of many recent political events: as evidenced especially in the social media (which have a known tendency to polarize judgment),[19] people are increasingly likely to treat those who adopt different social ideals as pariahs, as unworthy of moral regard. To put the matter crudely and in terms of the contemporary political situation in the United States, it is not merely that (cartoon) Red Staters have a different way of looking at the social world than (cartoon) Blue Staters. It is, rather, that at least some members of each of these two cohorts thinks of many members of the other cohort that they are not on an equal moral footing and are not worthy of moral regard and, indeed, are not the sorts of beings that you can bargain or deliberate with about jurisdictional rights or the benefits to each of the diversity of perspectives made available by the presence of the other.

And, indeed, Gaus is fully aware of this possibility and its implications, though he doesn't phrase the matter in the same way. He says (*TI,* 222):

[I]n the end, some perspectives will conclude that even the most fundamental elements of the Open Society are worse that no moral constitution at all. 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 to a 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, or some near approximation to it.… Such perspectives may live along with, but are not part of, the Open Society, treating

its rules as at best mere descriptive norms rather than moral injunctions.

The key question about this characterization is whether the *In*cluded Perspectives, and their adherents, are sufficiently numerous for it to be possible to form an Open Society even at the expense of excluding some other, "marginal" perspectives. It is not clear to me, as a purely empirical matter, what the relative dimensions now are in many of the so-called North Atlantic democracies of those individuals able and those unable to recognize that their ideals can't be insisted upon as a non-negotiable basis for the more broadly social ideals in their community. These, then, are the "extremities" of the playful heading of this section.

Of course, as Gaus reasonably remarks (*TI,* 222), "[t]here is no reason why we should leave ourselves at the mercy of those who refuse to live on terms that others can endorse." As a key task, the new program has, I believe, to consider (let this be an injunction of its positive heuristic) first of all how we can assess the relative balance of power, in our societies, between diversity-respecting and diversity-abhorring cohorts and, secondly, should the diversity-respecting be "outnumbered" by the diversity-abhorring, what we are, practically speaking, to do about that. There's no reason to suppose that these questions are going to be easy to answer. There is less reason to believe that even an answer will delineate a feasible program of social reform that will, if they are currently lacking, restore the preconditions for the open society. But the questions and issues are primarily empirical ones, both about the sociology and about the social psychology of diverse societies and this suggests a final positive heuristic injunction: We need to continue and extend the work of Cass Sunstein and others on the conditions under which such extremism can flourish. It, rather than bare diversity, may well represent the biggest challenge to the new program for political philosophy, and taking up this challenge will require expanding the domain of political philosophy to include other disciplines.

I called this paper "How can we do political philosophy?" With every new publication, Gaus is showing us how to do political philosophy in a way that is multi-disciplinary, diversity-respectful, and practically oriented. Most importantly, he has loosened the ties in his own work to established paradigms, partly by the rehabilitation of those left behind by the canonization of those paradigms (Popper on this occasion, P. F. Strawson in *The Order of Public Reason*), and partly by extending the tool-kit of concepts and modes of analysis that are drawn on in his work (as, for example,

with his appropriation of game-theoretical modelling). In my view, this work nails the coffin shut on the project of ideal theory. And it delineates a rich field of enquiry with numerous open questions, characterized by me in terms of the injunctions they set up as part of the positive heuristic for the new program. In particular, the new program needs to consider extremism, polarisation, demonization of "the Other", and not just the "tamer" forms of diversity. Otherwise, it risks making the same misstep as, I think, Rawls made when he tried to control diversity, first by means of normalization, and then by limiting it to those differences that arise between "reasonable" conceptions. While some idealisation is unavoidable even if we are not practicing ideal theory, we can't abstract away from extremism if we hope for the new program to be morally relevant to our current situation. That is a challenge worthy of the times, and it is one, I am happy to report, that Gaus is fully alive to, as shown, for instance, in his recent article "The Open Society and Its Friends", online in *The Critique,*[20] which is especially valuable, in my view, because, unlike a lot of horrified commentary on recent political events (though Gaus too recognizes the dangers), Gaus's analysis foregoes the complacent comforts of marginalizing "the Other". That, surely, is the counsel of wisdom. If, as the engineering metaphor would suggest, we are to take seriously the problem of getting from where we are to someplace better, we'd better be reasonably "it is what it is" about where we are. Otherwise, we are engaged in fantasizing not just the end-point, but the starting-point on "The Road to Utopia". Gaus takes the path less (frequently) chosen, and that might yet make all the difference.

NOTES

1   Just as we speak, in scientific methodology, of the essential tension between the conditions that facilitate innovation and those that support tradition, so too might we speak, or so I have argued (D'Agostino 1996), of a tension, unavoidable if not "essential", between remaining in theoretical contact with empirical contingencies, on the one hand, and providing some critical/normative leverage over those contingencies, on the other hand. This tension is, I believe, intrinsic to any political philosophy, including, on one interpretation anyway, ideal theory.

2   See for example Sunstein 2009.

3   According to Google Scholar, at 18 January 2017, this book has been cited nearly 64,000 times.

4   See, for instance, Gaus, 2011, III.7.3.

5   Muldoon, op. cit., secs. 2.7-8. See also Gaus, *TI,* II.4.1.

6   Muldoon is particularly good on this point. See op. cit., p. 118.

7   There are faint echoes here of Rawls's notion of pure proceduralism (Rawls 1971, p. 85).

8   See Weick 1976.

9   See for instance Epstein 2006.

10   An important intermediate layer is that of individual social agents, who will themselves have their own modes of evaluation and will experience in their own persons the effects of changes to social institutional arrangements.

11   On these matters generally, see D'Agostino 2003.

12   As Gaus documents (*TI,* 87), Karl Popper foresaw the extravagance of the conditions required if this notion of ideal theory were not to be empty. "This line of analysis led Popper to conclude that 'the Utopian approach can be saved only by the Platonic belief in one absolute and unchanging ideal, together with two further assumptions, namely (*a*) that there are rational methods to determine once and for all what the ideal is and (*b*) what the best means of its realization are. Only such far-reaching assumptions could prevent us from declaring the Utopian methodology to be utterly futile.'" And utterly futile it surely is, and is surely shown to be by Gaus's careful analysis.

13   The matter of confidence is meant to signal Gaus's excellent discussion of "The dilemmas of diversity" at *TI,* III.2.

14   See for instance Peters 2012.

15   See for instance Brennan 2010.

16   See for instance Jost and Major 2001.

17   In case (a) we can expect that individuals will mutually enhance each others' always probably only partially shared understanding of the social ideal. In case (b), we can hope for the development, among individuals who might otherwise be at odds with one another, of some fellow-feeling that may, in turn, support their mutual recognition.

18   See in particular Hayek, 1973, ch. 2.

19   See Sunstein 2006.

20   http://www.thecritique.com/articles/open-society-and-its-friends/, accessed 27 January 2017.

REFERENCES

Bauman, Z. (1987). *Legislators and Interpreters*. Cambridge: Polity Press.

Brennan, G. (2010). PPE: An Institutional View. *Politics, Philosophy & Economics* vol. 9(4).

D'Agostino, F. (1996). *Free Public Reason*. New York: Oxford University Press.

— (2003). *Incommensurability and Commensuration*. Aldershot: Ashgate.

—(2010). *Naturizing Epistemology*. Houndsmill: Palgrave Macmillan.

Epstein, J. (2006). *Generative Social Science*. Princeton: Princeton University Press,

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

—(2016). *The Tyranny of the Ideal*. Princeton & Oxford: Princeton University Press.

Hayek, F. (1973). *Law, Legislation and Liberty*. Vol. 1. London: Routledge & Kegan Paul.

Jost, J. and Major, B. (eds.) (2001). *The Psychology of Legitimacy*. Cambridge: Cambridge University Press.

Kuhn, T. (1970). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Lakatos, I. (1970). Falsification and the Methodology of Scientific Research Programmes. In: *Criticism and the Growth of Knowledge*, eds. Imre Lakatos & Alan Musgrave. Cambridge: Cambridge University Press.

Merton, R.(1936). The Unanticipated Consequences of Purposive Social Action. *American Sociological Review* vol. 1.6.

Millgram, E. (2015). *The Great Endarkenment*. Oxford & New York: Oxford University Press.

Muldoon, R. (2016). *Social Contract Theory for a Diverse World: Beyond Tolerance*. London & New York: Routledge.

Müller, J. (forthcoming). *Capitalizing on Political Disagreement: the Case for Polycentric Democracy*. London & New York: Routledge.

Peters, G. B. (2012). *Institutional Theory in Political Science*. 3rd edn., New York & London: Continuum.

Popper, K. (1962). *The Open Society and Its Enemies*. Vol. 1. London: Routledge & Kegan Paul.

Rawls, J. (1971). *A Theory of Justice*. Oxford: Oxford University Press.

Schoelandt, C. van (2015). Rawlsian Functionalism and the Problem of Coordination. Paper delivered to the Pacific Division of the American Philosophical Association.

Sunstein, C. (2006). *Infotopia*. Oxford & New York: Oxford University Press.

— (2009). *Going to Extremes*. Oxford: Oxford University Press, 2009.

Vallier, K. (2014). *Liberal Politics and Public Faith*. London & New York: Routledge.

Weick, K. (1976). Educational Organizations as Loosely Coupled Systems. *Administrative Science Quarterly*, vol. 21.1