
Political Philosophy as the Study of Complex Normative Systems

GERALD GAUS

Email: jerrygaus@gmail.com
Web: <http://www.gaus.biz>

It is an honor and a treat when innovative social theorists and philosophers take time out of their important work to think about one's own. The set of papers being published in COSMOS + TAXIS are especially flattering. Fred D'Agostino, Blain Neufeld, Scott E. Page, Kevin Vallier, Lori Watson and David Wiens all constructively engage *The Tyranny of the Ideal* and open up new issues to be explored. I am especially grateful to Ryan Muldoon for organizing this symposium (and, I should say, for all his work from which I have learned so much). I am also delighted that the symposium appears in COSMOS + TAXIS because the main theme of the book, which I hope to emphasize here (and which D'Agostino and Page bring out wonderfully in their essays) is that the subject matter of social philosophy is complex systems, something Hayek (1964 [2014], chap. 9) was one of the first to stress. Until, as Page puts it, "the imperative of complexity" is appreciated by political philosophers their work will remain what Hayek warned against—constructivist fantasies in which the critical problems of evaluative diversity, path-dependency, uncertainty, and interconnectivity are assumed away.

My comments focus on three main themes. I begin (Sec I) by taking up some questions of method. To a large extent *Tyranny* (as I shall call it, hopefully sans unfortunate self-reference) is adamant that political philosophy greatly benefits from the rigor of more formal ways of thinking. This, unfortunately, is one of the features of the book which causes the most resistance, as political philosophers are generally deeply averse to abstract model thinking, and often dismissive of "metaphors." I then turn (Sec II) to discussing some aspects of the rugged landscape model I employ, and what I think it tells us about the nature of normative thinking. Section III turns to what D'Agostino calls my "constructive intention," the account of the Open Society. Many issues arise here, of which only a few can be addressed. In these remarks I make no effort to respond to all the ideas and, yes, criticisms, in these thoughtful essays, but I do believe that many fundamental points will be addressed—

hopefully in a way that enlightens readers who are not especially concerned whether *Tyranny* is bullet proof (it is not).

I. MODELS

Abstractness and Modeling

An explicit aim of *Tyranny* is to model a long-standing problem in political philosophy in a more rigorous and abstract way, which I believe alerts us to critical features of political theorizing that have hitherto gone unnoticed. When we model a phenomenon we always abstract from some of its features to better understand the working of others. All modeling is a selection process. We construct a simplified, abstracted, analysis to get insight into critical features that are obscured by more detailed descriptions. It does not follow that the features from which we abstract are unimportant; in another context we may construct a model to better understand *them*, perhaps putting aside the very features our first model highlighted. As Michael Weisberg (2013, p. 13), points out, philosophers generally prefer models that are rich in description, and this perhaps leads them to so often object to a model because it has left something out or is "blind" to it (as, for example, *Tyranny's* model quite intentionally abstracts from the "logic" of justification advanced by a particular perspective—more on that anon). The aim of *Tyranny's* modeling approach is to abstract from many features that have traditionally been in the foreground to discover those thus far overlooked. And the feature that I wished to stress—which I tried to show is implicit in the presuppositions of many "ideal theories"—is that seeking ideal justice involves optimization in a certain type of structured complex system. Once we understand that ideal theories are committed to analyzing the pursuit of justice as an optimization exercise in, strictly speaking, a mathematically complex system, we see them and their problems in an entirely different light. Indeed, we see political philosophy in a new light. I believe that this rather startling insight justifies abstracting from so many of the features and issues that have hitherto been the concerns of political philosophy.

In a way my intention was to present political philosophers with an unfamiliar type of analysis, for it is through confrontation with the alien that we appreciate our presuppositions. That, indeed, is the theme of the entire book.

Tyranny advances an analytic result. If ideals are characterized by institutions, and institutions are “coupled” and so have “interactions” or “interdependencies” in their resulting justice, then as Page nicely summarizes in his essay, under the conditions I specify we are almost certainly confronted with “The Choice” between pursuit of local optima (local improvements in justice) and the global optimum (ideal justice). This is an inescapable conclusion. Of course one can dispute the assumptions or the applicability of the model, but not the well-nigh inevitability of The Choice given them. To be sure, as Neufeld and Watson note, this makes the argument abstract (given the norm in political philosophy), but I believe that is a good thing, for it helps us to achieve distance from our ideological convictions and intuitions to see the relations between our commitments in ways that may surprise and enlighten us (Johnson 2014). And as we abstract from the details, we open up ourselves to what Page calls “transdisciplinary” insights—that the logic of the problem in, say, politics, is similar to problems in management or evolutionary theory (Lane 2017, chap. 1). Thus *Tyranny* argues that Kauffman’s *NK* and the Hong-Page models provide insights into what we might have thought was a distinctive feature of “ideal political theory.” To be sure, we must proceed with care to ensure that our model captures the fundamental features we are interested in—thus my significant modifications of the Hong-Page model in understanding ideal political philosophy. Still, the aim is to abstract and see if we can distill the problem to its essentials, and so provide a general result. So like Rawls (2005, p. lx), I do not apologize for the abstract nature of the analysis.

Of Metaphors and Models

As Page explains in his essay (and in his path-breaking and, if I may say, often stunningly innovative work over the last two decades) different models can be employed to understand complex systems. I focused on what have been deemed “rugged landscape models.” My idea was this. Some political philosophers have hit upon the metaphor of mountain climbing with two core dimensions: climbing up (achieving more justice) and moving laterally (getting closer to the arrangement that characterizes the most just social state, the ideal). The thought is that to move closer to the ideal on the lateral dimension may sometimes require first moving up

(getting more just) but then, like an intrepid climber, going down for a while before forging up the next slope. Thus at times we are decreasing justice (altitude) in order to make our society more like the ideal (a closer latitude). (We can make the model more complex by adding other dimensions (*Tyranny*, pp. 258-9) but two were enough for my purposes).

Wiens’s essay focuses on metaphors and models: what are we doing when we employ “metaphors,” “models” (and “theories”)? There is no canonical view; in her classic work Mary B. Hesse (1966) argued that models just *are* metaphors. I offer a somewhat different account here, according to which metaphors are typically basic or initial models. In scientific explanation we can distinguish primary and secondary systems; the primary system is the phenomenon (*A*) to be explained, the secondary system is the explanation of *A* in terms of the metaphor/model (*B*) (Hess 1966, p. 158). “Sound (primary system) is propagated by wave motion (secondary system)” (Hess 1966, pp. 158-9). Metaphors are critical in suggesting analogies, such that the primary system can be understood in terms of the working of some other system—which it patently is not. “For the conjunction of terms drawn from the primary and secondary systems to constitute a metaphor it is necessary that there should be patent falsehood or even absurdity in taking the conjunction literally. Man is not, literally, a wolf, gases are not in the usual sense collections of massive particles” (Hess 1966, p. 160). To take, say, navigating a mountain range (*B*) as a metaphor for pursuit of the ideal social state is to suggest that we can begin to understand the primary system *A* (pursuing the ideal) as having similar relations to navigating a mountain range (*B*). It is important, as Hess notes, that metaphors are not mere similes: we do not know ahead of time in just what way *A* is like *B*, or the ways in which it functions analogously. When we think about *A* as a *B*, we begin to think about *A* as acting as we know *B* does, and we look for familiar features of *B* to see if they are in *A* as well. We can use *B* as the basis of analogies that help us understand some of the puzzling workings of *A*, and provide the basis of further investigations.

This already is a model; we focus on the features of the primary system that are revealed by our more thorough grasp of the metaphorical secondary system. At this stage *B*’s dynamics and features are critical in picking out and understanding *A*’s; still employing analogies we search for parts of *A* that seem a lot like *B*’s, and use our knowledge of the way *B* works with those parts to model how *A* must work too. For example, early electricians sought to understand electrical phenomena (*A*) as a liquid (*B*); and since *B*

could be bottled, so should we be able to bottle *A*—hence their successful effort to develop the Leyden jar (Kuhn 1970, p. 17). In my view this metaphor-analogy stage is often the first step in building a rigorous model of *A* that ultimately jettisons the metaphorical *B* as the base comparison and, so, no longer employs metaphors or analogies. The model is stated in axioms and equations, and variations in these, and anomalies encountered by our current model, suggest further developments.

Perhaps even after a formal model is developed the original metaphor may be used as a basis for speculating what variations of the formal model are worth exploring, though often (and this is my point) at some juncture the original metaphorical base, *B*, may become a hindrance rather than a source of further discovery, even if we continue to keep some of the labels suggested by the now-discarded metaphor. Thinking of conflict as a game blossomed in Prussia in the nineteenth century in the fad of “Kriegspiel” (literally, “war game”)—a board game of conflict, played both by the public and the general staff. Indeed, in 1825 the German chief of staff proclaimed “It is not a game at all. It’s training for war!” (Poundstone 1992, chap. 3). Young John von Neumann and his brothers played their own version of Kriegspiel. Game theory built on this metaphor; and we still have terms like “players,” “moves,” “strategies,” and so on. But von Neumann and others developed “game” theory such that now its categories and relations are strictly defined within it, and so reference back to the way board games function is not apropos. It has its own, well-defined, concepts and mechanisms—in Wiens’ terms its own “mathematical objects”—and any lingering game terms are simply for ease of exposition or to help beginners by invoking in their minds the now-discarded metaphor, which can help them to begin to see their way around the fully formalized model. (On the other hand, the atavistic labels can be an impediment, as when neophytes are told that they are “playing a game” and so, going back to a “game frame,” play iterated prisoner’s dilemmas not to maximize their own outcomes but to ensure that they “beat” the other “player”). In my view, however, game theory is not a metaphor: it is a formal model. We explain the primary system (people in interdependent actions) in terms of a formalized secondary system.

Hence my basic idea. The mountain range metaphor recurring in the political philosophy of ideal justice is a basic model that gives us a clue to some important features, but this metaphor has been superseded by fully formalized and developed models employed in fields such as evolution-

ary biology and complexity theory. The development from model-as-metaphor to a fully axiomatized (which is not to say fully developed) model has already been made, and was awaiting exploitation by political philosophers. As Page explains in his essay, I relied extensively on Stuart Kauffman’s *NK* model. In Kauffman’s model, all the terms are fully specified and the relations mathematically determined. There are Boolean nodes (*N*) that may be linked to *K* other nodes each of which can be “turned on or off” by their connected nodes. If we are modelling genes, the state of each node (gene) affects the fitness of the organism. When *K* is greater than 0, some genes are interconnected, and so the overall fitness of the organism will not simply be an additive function of the fitness of each gene, but of their number and degree of interaction.¹ This implies that when $K=0$, an organism *O'* that is a one gene variant of *O* will have a fitness highly correlated with the fitness of *O*. When $K=N-1$, the fitness of *O'* will not be correlated with *O*. For ease of exposition we can label the former a “smooth optimization landscape” (when we graph the fitness of variants they will increase or decrease in smooth lines) and the latter “maximally rugged” (fitness values may jump all over the range from one variant to the next). On all but the smoothest such optimization “landscapes” there will be local optima (any one gene variant will be less fit), which are not the global optimum. Each optima can be called a “peak.” For now the critical point is that once the model is specified “landscape,” “ruggedness,” “peaks,” “height” and so on are fully determined by the value function (e.g., adaptiveness, justice), the structure of the domain (the measure of genetic or world variations) and the *NK* dynamics. These terms are no more metaphorical than “player” and “strategy” are in game theory; they can be replaced by purely formal notation and nothing would change.

I have spent perhaps too much space on the relation of metaphors and models, but I have repeatedly encountered political philosophers who, confronting the rugged landscape model, conclude “well, that’s a nice (or bad) metaphor, but let’s get beyond metaphorical talk and do some *real* political philosophy” (e.g., contemplate our intuitions or the perennially-fascinating question of whether “ought” implies “can”). I hope it is clear that this is a basic misunderstanding of the place of both metaphors and formalized models in inquiry.

II. COMPLEXITY AND OPTIMIZATION

Pursuit of an Ideal and Ruggedness

Let us build on Page's {Market, Bureaucracy, Democracy} model to better see how the complex optimization model works. He writes:

In the model, a society must allocate resources and opportunities across a set of domains. Within each domain, the society chooses among three pure institutional types: a market (M), a bureaucratic organization (B), or a democratic mechanism (D). For example, to select a construction firm to build roads, a society could hold an auction among qualified firms (M), it could construct a bureaucracy that develops criteria for selecting a firm (B), or it could hold a vote among elected representatives for the winner (D). If there exist ten such domains, then the set {X} consists of all vectors of length ten whose entries belong to the set {M,B,D}. Though a simplified characterization of the world, the model allows for a combinatorial explosion of social arrangements—59,049 distinct possibilities to be precise.

We see how quickly the number of distinct possible social worlds expand. To simplify, suppose we have only a four-domain world, any of which can be organized on market, bureaucratic or democratic institutions. Suppose:

- Domain 1: Decisions about supply of public goods
- Domain 2: Decisions about the supply and distribution of private goods
- Domain 3: Decisions about income distribution
- Domain 4: Decisions about the distribution of employment

Suppose we have a certain market socialist ideal, {DMDD} (public goods, democratic; private goods, market; income distribution, democratic; employment allocation, democratic). We are now at {BMMM}. One thing we might do is simply list the justice of all 81 possibilities, but it is hard to know, say, how the {MBDM} world would function: what would a social world be like where public goods are determined by the market, private goods are distributed by a bureaucracy, income distribution is democratically voted upon, but allocation of employment is via the market? Hmm. The ways all these mechanisms would interact are,

Tyranny claims, extraordinarily difficult for us to model and predict, as our social science is based on understanding market provision of private goods and market (and some bureaucratic) distribution of incomes.

On *Tyranny's* analysis, ideal theory presents a perspective on justice that *orients* our quest for perfect justice by locating the ideal in relation to our current social state. A perspective (i) identifies what social states are similar to others (critical to orienting us in the quest for justice), and (ii) assigns a justice score to social states given their expected functioning. Considering just the first function, in our case we might have, say, the following perspective:

{BMBM}—{BMMM}—{DMMM}—{DMBM}—{DMDD}
(ideal)

On this perspective we are presently at {BMMM}; public goods are allocated by a bureaucracy, all others by the market. According to this democratic socialist perspective the ideal is a condition where public goods are allocated democratically, private goods by the market, while incomes and employment opportunities are decided democratically {DMDD}. Now on this perspective a world where public goods are decided by a democratic rather than a bureaucratic mechanism—{DMMM}—is pretty close to ours, and we can estimate how it would function and its justice. After all, for a democratic vote to replace a bureaucratic decision does not seem a huge jump in the social space. Bringing in now the second (value) function of a perspective, suppose this perspective judges that we are presently at 50, and the ideal is 100. What about {DMMM}? Here is a distinct possibility: while moving from {BMMM} to {DMMM} makes our society's structure closer to the ideal {DMDD}, it could decrease justice. In {DMMM} public goods are decided by vote but incomes by the market. Many social democrats such as Rawls believe that great wealth corrupts democracy, and so the provision of public goods in {DMMM} might produce a less just society, where the wealthy control their provision more than at present (so justice goes down to 40). On the other hand {BMBM}, instead of democratizing decisions about income distribution, puts them under bureaucratic decision-making. This, plausibly, leads us away from "bourgeois democracy" and its ideal of self-government by instituting an expert bureaucratic elite who end up exercising great control over the economy (Schumpeter 1950, 296ff; Levy and Peart 2017). Yet, it may lead to more distributive justice, and perhaps is superior in overall socialist justice to {DMMM}, say 60. Hence "The Choice" in Figure 1.

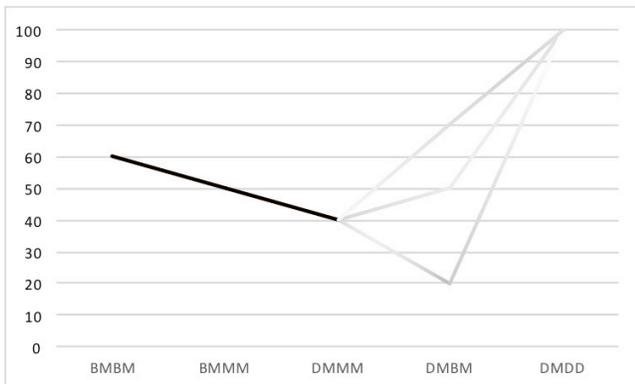


Figure 1.

At our present {BMMM} we have a choice between moving to {BMBM} with a justice of 60, but at the cost of making our society’s institutional structure less like the ideal (we move away from it). If we seek to make our society conform more closely to the institutional structure of the ideal, we will suffer a loss in justice. After that (as indicated by the multiple light grey lines), we are uncertain as to the justice of the next step, as we are modeling a world where the bureaucracy controls income distribution but not employment allocation. The ideal theorist need not always choose to move toward the ideal, but if the ideal theory is to provide significant guidance, she must often choose to pursue the ideal and forgo local improvements in justice. The Choice is really a choice.

It might be wondered, if we know so little about {DMBM}, how do we know so much about {DMDD}, the ideal? And, of course, that is *the* question: how can we be sure of the functioning, and so justice, of a set of social arrangements that are so different from our own? Should we ever get to {DMBM}, it is almost certain that we will drastically re-evaluate the ideal, {DMDD}. As we approach the ideal, it changes before our eyes, perhaps receding into the distance (or perhaps, alas, we spot it in our rear-view mirror).

As the number of domains, dimensions of evaluation, or distinct institutional structures involved in an ideally just society increases (N), and as the justice-relevant interdependencies among them increases (K), it becomes essentially certain that The Choice will arise. Nuefeld and Watson are certainly right that institutions must be “mutually realizable” but that is not enough for their evaluation given the inevitable coupling in their functionings: some mutual realizations lead to excellent interactions (from the perspective of justice) while others interact in detrimental ways. This is the root of the complexity and uncertainty that confronts

all attempts to move us from one institutional scheme to a very different one.

I think it is important to stress that this is not a just-so modeling story. In their extensive fieldwork on real institutions Elinor and Vincent Ostrom stressed that institutions are composed of numerous rule configurations; the constituent rules have strong interdependencies, both with each other and with environmental conditions. “A change in any one of these variables produces a different action situation and may lead to very different outcomes” (E. Ostrom 1986 [2014], p. 111). When we talk about social states such as a “property-owning democracy” we are referring in a loose way to a myriad of interconnected rules and behavioral tendencies that constitute the working of the set of institutions that are summed up by this moniker. Building these institutions inevitably leads to problems of complex optimization and searching.

The Fundamental Diversity Insight

Vincent Ostrom (1972 [1999], p. 125) once remarked in respect to his empirical findings, “[t]he complexity of relationships ... is such that mortal human beings can never observe the ‘whole picture.’ Anyone who attempts to ‘see’ the ‘whole picture’ will ‘see’ only what is in the ‘eyes’ or ‘mind’ of the beholder.” This insight is formalized in the Hong-Page model, where diverse perspectives, each seeing (coding) the problem in different ways, each see possibilities to which the other is “blind.” For example, one might wonder about Figure 1: why is BMBM a move *away* from the ideal, rather than step toward it? The Choice would not be confronted by the perspective:

$$\{DMMM\} - \{\mathbf{BMMM}\} - \{BMBM\} - \{DMBM\} - \{DMDD\}$$

(ideal)

The Hong-Page model shows us how different perspectives on a problem can help each other in their searches for the best outcome. We can benefit from interactions with those who see the world differently: what is a tricky problem for me might, given your perspective, be an easy one. This is a fundamental insight, and *Tyranny* spends a good deal of time evaluating its applicability to the search for ideal justice. Page’s essay nicely summarizes both the Hong-Page model’s resources, and the reasons why I conclude that it has a restricted applicability to the problem of ideal theory. For now, I merely stress that I strongly endorse (stated approximately)

The Fundamental Diversity Insight: Any given perspective Σ on ideal justice is apt to get stuck at poor local optima; other perspectives can help by reinterpreting the problem or applying different predictive models, showing better alternatives to be in Σ 's present neighborhood.

Wiens believes the arguments for this and related diversity-relevant conclusions do not “rely much” on the rugged landscape (*NK*) analysis. I do not agree. It is just because we are navigating this sort of problem that our perspectives on justice are so apt to get stuck on poor local optima, from which we confront The Choice. Critically, as the Hong-Page model taught us, other perspectives can—sometimes—point the way forward.

A Tale of Two Models

Wiens's essay is, to a large extent, a contrast between his optimization model and the model we have been examining. Trying to be a little less formal, Wiens's (as I shall call it) “simple optimization model” seems characterized by:

1. A set of possibilities, X (options, states of affairs, social worlds, etc.). This is the domain of the value function on which it operates. In *Tyranny* they are alternative social worlds.
2. A binary value function (which I'll call V): the value relations among members of X are built up through binary comparisons. Letting (x, y, z) be elements of X (i.e., particular social worlds), xRy if and only if x is ranked by V at least as high as y . If xRy and yRx , x and y are ranked the same; if xRy and not yRx , x is ranked above y .
3. It also seems required that the binary relations generated by V be transitive among all triples in X ; $(xRy) \& (yRz)$ implies xRz . Another requirement seems to be completeness, i.e., for all members of the domain X (x, y), either xRy and/or yRx . Other conditions on the ordering are allowed.
4. There is an additional set of constraints. These allow us to partition the domain, X , of possible social worlds, into those that meet these constraints and those that do not. Call X_C the subset of X that meets some given constraints C .

As I understand Wiens (and I'm sure I have not grasped all the intricacies of his model, so apologies in advance), conditions 1-4 imply that the only structure among the ele-

ments of X that are not based on V (the binary value relation) is the partition of X into the subset X_C and the rest of X . Consider, then, simply the X_C partition. For all the social worlds in X_C the only structure relating them is that yielded by V . Until the binary ordering is applied to the members of X_C , they are an unordered set of social worlds that meet certain constraints.

Perhaps the most fundamental difference between simple optimization and *NK*-optimization is that the latter models a structure among the members of the domain (possible social worlds) that is not generated by the value function. As Page notes in his essay, we can make this point by saying that the simple optimization model's theory of justice T only has a measure of the range of T (the ordering produced by the value function) while *Tyranny*'s model has independent metrics for the range and domain of T . Without these two distinct metrics, as we have seen, The Choice does not arise. Rather than confronting an unordered set of possibilities to which we apply our value (justice) function, *Tyranny* assumes that all the members of X —“social worlds”—have certain justice-relevant features and these generate a structured domain, to which we apply our value function. (We should refrain from calling this domain metric simply “descriptive,” as it is generated by the similarity of worlds' justice-relevant features, so they are, we might say, normatively loaded descriptions.) And in most of these landscape models—e.g., those in evolutionary biology—the structure of the option set is the similarity of the constituent features of the options, such as genotypes. If x is almost but not quite identical to y in the relevant respects (and, as always, “relevancy” is defined by the perspective or theory being employed, see below), then before any value function is applied, x is located close to y ; if the defining features of z share very little with x , x and z will be located far apart. In a simple evolutionary biology model, y might be a one gene variation from x , and so we can entirely correctly say—before we know their adaptiveness (value), x and y are close (though it could turn out that in terms of adaptiveness they are not).

What Model Should We Use?

So which model should we use? A model that only looks at the range of value (i.e., the justice of the options) or one that includes a meaningful structure (i.e., the institutional similarity) of the elements of the domain? As soon as we phrase the question so bluntly, we see how misguided it is. The Fundamental Diversity Insight indicates that different models often provide different insights, so we are all apt to

benefit when multiple models are pursued (Page 2016). The model we employ depends on our perspective on the primary system and our theoretical aims. As we know from economics, in many cases a simple optimization model, which only orders according to the value function (i.e., utility), elegantly analyzes the problem, say of most consumer choice. But as Harold Hotelling (1929) showed in his analysis of the location of shops, sometimes we get more insight by also including a structure of the option set—in Hotelling’s case, the geographical location of shops along a street. Anthony Downs (1956) saw that this “spatial” model could be extended to include “ideological space” (as well as utility values defined by the value function) and so began spatial models of politics. Downs’s model was not metaphorical, even though it was developed from a “geographical” model: ideological space was well-defined.

Hotelling’s and Downs’s models included a structured domain (as well as a value function), but were not *NK* models. I have tried to indicate why I believe the ideal of orienting our quest for justice by a fully just social state is well modeled by a rugged landscape in which our reliable knowledge is confined to our current neighborhood. Given the large number of institutions and background conditions that constitute an ideal social state, and the myriad of couplings between them that result in vastly different social states of different degrees of justice, setting out to pursue the ideal is an exercise in Knightian uncertainty (Knight, 1921 [1964], chaps. 7-8). For the most part, we only have useful assignments of probabilities within our neighborhood, so we simply do not have the information necessary for a fruitful simple optimization exercise.

Consider a mundane case: a manufacturer searching for innovations in product *P*. The advice to formulate an ordered list of possible variations of *P*, and then take the best choice is not of much help; the manufacturer cannot assign values to many of the options. Indeed, she doesn’t yet even have the blueprints for many variations. So here is an entirely sensible approach to the development problem. Have most of the research teams work on near improvements (slight modifications of the technology underlying *P*); because there is so little fiddling with the structure of the present *P* we’ll often find the new versions slightly better or slightly worse, and can further build on the slightly better ones, and then build on some of those, etc. This is a conservative “climb the gradient” heuristic, always seeking local improvements. But we may get stuck at a product for which no small improvements could be made (yet is not the ideal *P*). Think of SONY teams that were working on

the best Betamax recorder.² So the manufacturer may also wish to invest in an R&D department that has some teams working on more radical innovations (say, laser discs), some of which could result in really high values, but it is almost impossible at this point to make sound judgments about whether they will pan out. Here we are dealing with hunches, hopes and dreams—not probability assignments. In the case just described, we are searching the value of options with a certain structure, and these structures are the very properties that yield valuable products. This structured space thus *orients* the product search. It is important that in this case one’s optimization problem *starts* from a location (a place in the structure), and we are thinking of how to best move *given where we already are*.

The claim of *Tyranny* is that an interesting class of ideal theories (I never say “all”) have much in common with this case, except that these ideal theorists believe that they have already developed the blueprint for the perfect (or at least truly excellent) “product” but we can’t build it right now (the blueprint may require components not yet developed, like public-spirited folk). Our job is, given our present location in the domain, to begin to develop structures that are more like those that generate this great result. That is what it means to *build* an ideally just society. To build is to assemble the components. Because the task of such ideal theories is to work from our given structure to the ideal one (and thereby achieve the perfect value score), an optimization model that includes structure (which is not simply generated by the value function) is necessary.³ It is in this sense that, in a structured domain, the ideal “orients” improvements in justice in a much more complex sense than does the simple injunction “given some set of unordered social worlds, choose by maximizing the value function subject to constraints.”

Does a Simple Optimization Model Capture Ideal Theory?

Although I favor an ecumenical approaching to modeling, I am not a model nihilist who supposes that one modeling choice is as good as the next. *Tyranny* adopts a theoretical perspective on analyzing theories of ideal justice which, I think, is in many respects superior to a simple optimization model based on a binary value relation. Deriving “perfect” or even “good enough” from a value function that can only yield judgments about what is “better” is, I think, a job and a half. Theories based on a binary value function can only say that an option is the best in the sense that it is better than all the options in some set of options. So to say that a certain social state is ideally just is, essentially, to say that

it is better than all other social states in the domain X , or more realistically in some X_C partition. This makes it both too easy and too hard to find the ideal. It is too easy, because when a theory has some non-empty option set and can identify a best element, the theory then has apparently located an “ideal”—essentially every optimization exercise ends up as an ideal theory (which is why Wiens can see the Open Society as an ideal, since I think there is a set of “devices” that are maximal choices, even if no set is the best choice, to accommodate diversity). On the other hand, it is too hard to identify the ideal, as we need an exhaustive enumeration of all elements in the set X_C of possible worlds to identify a “best” that is not simply dependent on the choice of the comparison set.⁴ Thus we can only be confident that i is the ideal world if we are confident i is not merely the best in some subset of X_C , but remains best when the set comprises all of X_C : i is better than each and every element in X_C . In formal terms, it must be the unique choice set from X_C . Showing that is a daunting task, unless we very tightly specify the constraints that define X_C . I don’t think this is the usual way that ideal theorists have reasoned in the history of political thought. On my view, *The Republic*, More’s *Utopia* and Bacon’s *New Atlantis* are paradigmatic. Construct first an imaginary social world in which humans would relate to each other with full justice, etc., and *that* is the ideal. We score *it* as fully just, without a complete ordering of all worlds in X_C . That then becomes our inspiration, even if we do not know of many of the other worlds in X_C . Indeed, the common utopian theme that the ideal is a far-off land to which we have yet to find a path-to-construct-it implies that there are many unknown options between it and where we now are.

Of course to inspire it does have to be absolutely perfect—we might construct a world that is within the human horizon of workability and yet is essentially but not ideally just. For Rawls identifying such a “realistic utopia” was a critical task of political philosophy. At various times he agonizes over the question whether a reasonably just society is within the grasp of humans. “The wars of this century with their extreme violence and increasing destructiveness, culminating in the manic evil of the Holocaust, raise in an acute way the question whether political relations must be governed by power and coercion alone. If a reasonably just society that subordinates power to its aims is not possible and people are largely amoral, if not incurably cynical and self-centered, one might ask with Kant whether it is worthwhile for human beings to live on the earth” (Rawls 2005, p. lx). It is hard to see how showing some social world is

“ranked first in our option set” could possibly assuage such worries. Rawls seeks to find out whether the *best* among our options is *truly just*—and if there is one that is truly just it is to orient our long-term endeavors at reform.

Perspectives, Normalization and Generality

As we have seen following (though, as he points out, substantially modifying) Page’s framework, I argue that a *perspective* provides an orienting structure to the complex optimization problem. A perspective on justice includes, as it were, all the elements needed to generate an ideal theory. It includes a set of evaluative criteria (perhaps liberty and reciprocity, perhaps sanctity and respect for authority), an identification of what parts of a social world are relevant when evaluating it in the light of these criteria and a model of the way these features interact to provide an overall social order, which then can be scored in terms of its justice. In addition, a perspective must have some view of how similar worlds are given their justice-relevant features; it must be able to say world x is almost identical to y in terms of the institutions, rules, motivations and so on that define the worlds. Now fundamental to *Tyranny* is that how all this is accomplished is *internal to a perspective*: in Wiens’s language these *must* be “black boxes” the contents of which the model is “blind” to. Political philosophy is normally devoted to explicating the correct perspective, giving accounts of what the correct evaluative criteria are, how they should be combined and what institutions are relevant to justice. As D’Agostino puts it, this is part and parcel of the “legislative” stance in political theorizing. In addition, an ideal political philosophy seeks to show what the ideal is, and how far we are from it. Call this a *fully normalized view of ideal justice*. Sometimes, as with social contract theory, a set of perspectives is identified as “correct enough” because of their similarities in evaluative standards or their basic agreement on the relevant institutions; call this a *partially normalized view of ideal justice*. It was Rawls (2007, p. 226) who notes that all social contract theory supposes some normalization—all take a peek inside the box and seek to give us some idea how it ought to function.

Of course we often do care about exploring evaluative criteria or, for Wiens, the logic of justification. When we do so we seek to develop a fully or partly normalized theory—one that commences by identifying a set of correct or minimally acceptable perspectives. However, the aim of *Tyranny* was to present a fully general model of ideal theory, one which analyzed ideal theory *qua* ideal theory, and not *qua* liberal ideal theory, or *qua* ideal theories that embrace a view of

the logic of justification, or even ideal theories that accept a standard secular understanding of social reality. Two aims were regulating here. First, as I have been stressing, the aim was a fully general theory that uncovered the logic of ideal theorizing itself, as far as possible freed from familiar substantive commitments (a “pure theory of the ideal,” we might say) and, secondly, to explore how a society with a truly radical diversity of perspectives, each committed to its own theory of justice, might not only live together, but learn from each other. It was thus important not to start out by imposing any normalization on perspectives.

Still, as Wiens perhaps suggests, isn't the model itself a form of normalization? It says that a perspective must have evaluative criteria, must identify features of the social world, must be able to score them and so on. Isn't this just another normalization? We might call this the “all theory implies normalization” objection. And this raises one of the most perplexing issues for what we might call ecumenical theories of a phenomenon (see Gaus, 2017b). Think of the difference between an objective theory of value (say, the labor theory) and a subjective one (such as Carl Menger's). On the one hand, a subjective theory seeks to allow that value is, in a basic sense, up to the agent, and she can place value where she wills. In that sense it is ecumenical about what can be “correctly valued.” Yet, because it is a theory of value it must delimit its inquiry—valuing isn't the same thing as sleeping (even if sleeping is valuable). To give a theory of *X*, one must identify a class of *X*-phenomenon. Unless we characterize *X* in *some* way we can't even begin to talk about “it,” and unless our characterization has enough structure we can't begin to analyze it. So, yes, categorizations and specifications cannot be avoided in theorizing. If all categorization implies normalization, so be it. The important point for *Tyranny* is that, having identified the basic elements of a perspective, the analysis does not proceed to employ additional criteria to distinguish “good” v. “bad” or “reasonable” v. “unreasonable” variants (though each perspective may make such judgments of the others, as we are about to see).

III. THE OPEN SOCIETY

The Fundamental Diversity Dilemma

As I said above, *Tyranny* relies on Hong-Page reasoning to endorse *The Fundamental Diversity Insight*. A point of departure from their model is my analysis of, and emphasis on, the *Fundamental Diversity Dilemma*. A radically different perspective from Σ may provide important insights to

Σ , but adherents of Σ will have great difficulty making sense of this radically different perspective. As Σ sees them, they categorize the social world in very odd ways, and employ strange evaluative standards. This should be familiar to everyone. Think of the radical atheist's claim that religious perspectives suffer from a cognitive dementia, or the religious view that the atheist is infected by evil. Under these conditions both see the other as saying barely intelligible things, and they certainly do not find each other's “searches” for the ideal of value. In a diverse society, sometimes we can learn a lot from other perspectives, sometimes we can use some of their discoveries, and sometimes we will dismiss them as lunatics.

As D'Agostino has argued in his marvelous *Naturalizing Epistemology* (2010), the trick is to “get it together,” to assemble the insights of our different research programs in a way that leads to mutual enlightenment. Adapting D'Agostino's analysis of scientific communities, I employed his idea of a “republican community” to designate a community of moral inquiry, i.e., one that shares sufficient standards, problems and concerns such that the results of the searches of some in the community can be taken up by others. Of course such “republican communities” come in all varieties, from the thick who share many deep commitments, to those who sometimes find some results of the others as somewhat enlightening. Employing a “small world model” *Tyranny* argues that a society composed of diverse groups, some of whom entirely dismiss each other's insights, might, nevertheless, be one in which everyone learns from everyone else. The basic idea is straightforward. Our Radical Atheist proposes committing the Evangelical Christian to an asylum, while the Evangelical Christian dismisses the Atheist as vigorously. They will never directly share insights. But, say, a Roman Catholic Scientist may engage the Evangelical and, in turn, what we might call the Broadminded Secular Scientist is willing to engage the Roman Catholic Scientist; so there is a line of engagement from the Evangelical Christian to the Broadminded Secular Scientist. If the Radical Atheist engages the Broadminded Scientist, the network of mutual influence is complete. The Evangelical Christian and the Radical Atheist may end up enlightening each other. As we multiply the number of perspectives, diverse, crisscrossing networks of various “republican communities” of inquiry will arise. Such networks mitigate the Fundamental Diversity Dilemma. The Open Society, I argue, provides a framework that allows these networks to develop, and so a framework that allows each view of justice to better cope with its own internal challenges.

Communities of Moral Inquiry v. Moral Communities

Vallier's essay raises the important question of the relation of "republican" communities of inquiry to moral rule networks—people who share moral expectations and demands. Let us call a "moral community" a set of individuals who (i) share very similar perspectives on justice, (ii) interact on moral rules grounded on these perspectives and (iii) share very few moral rules with those outside of (ii). In the history of political philosophy, many have thought that an ideal way to cope with moral diversity would be for us to divide up into numerous like-minded moral communities (Nozick 1974, chap. 10). *Tyranny* tries to show why that is an error. In a network in which the moral rules reflect only similar perspectives, members are able to act on views of justice they take as superior, but at the same time they lose resources (interaction with other perspectives) that have the real potential to enlighten them about justice. What constitutes a "republican community of inquiry" is in constant flux; as Σ develops, perspectives that Σ previously saw as beyond the pale of sensible inquiry become intelligible, while some that it previously deemed sensible may now appear antiquated. *Tyranny's* concern is with those who are devoted to knowing what real justice is, and that cannot be secured in like-minded moral communities. Thus the crux of *Tyranny*: the subjects of my inquiry—those who are deeply devoted to knowing what a just society is—have powerful reasons to embrace the diversity of the Open Society.

One important difference between *The Order of Public Reason* (2011) and *Tyranny*, then, is that the latter is focused on those who are convinced there is a notion of perfect justice, and are committed to knowing what it is. Such folks would seem to be an especially hard case for a theory of public reason as given in *The Order of Public Reason*, as their primary concern is *getting justice right*. Why would they, of all people, want to share a moral order with people who are getting it *wrong*, much less accommodate them in some way? *Tyranny* (the book, that is), is an extended answer. Surprisingly (at least it was surprising to me), if you want to really get justice right, you must live with, and learn from, those who get it wrong—if knowing the perfectly just society is a complex problem.

Polycentrism and Accountability

A second difference concerns the model of moral relations. As Vallier points out, *The Order of Public Reason* analyzed morality in terms of the social rules endorsed in some group, *G*. It was supposed that the persons in *G* share a network of moral rules. As I noted, groups come in many dif-

ferent sizes, but my concern was typically with the largest interacting group, what Hayek called "the Great Society." *Tyranny* develops a more nuanced account—polycentric moral networks. What I have called "moral communities" are typically confused with what Elinor and Vincent Ostrom called "polycentrism." The Ostroms argued that a "highly fragmented" system in which different groups were largely confined to their own "jurisdictions" (or communities) is apt to result in conflict and institutional failure (Ostrom and Ostrom 1977 [1999], p. 96). They thus modeled a successful polycentric order in terms of many crisscrossing and overlapping jurisdictions and norm networks. *Tyranny* relies on polycentrism thus properly understood. In *Tyranny* "the group" dissolves into complex and overlapping moral rule networks. On some fundamental matters, the moral rule network is coextensive with the Great Society of strangers, the group that I had in mind in *The Order of Public Reason*; but on many other matters "the group" disaggregates into moral networks that are subsets of *G*. As the Ostroms stressed, whether moral rules and institutions are needed in the eyes of some individuals depends on what sort of problem they are facing; some problems can only be solved with the participation of almost everyone, others have much more restricted scope.

Within each moral network—i.e., people who have shared normative and empirical expectations about what is to be done in some circumstances—each is accountable to others in his network for violating the rule.⁵ This is not the case in relation to outsiders: one is not answerable to outsiders for one's failure to conform, nor can they hold one accountable. So if, in a moral network of vegetarians, Alf defects and has *foie gras*, he is not answerable to me (a devoted carnivore), unless we begin to tell a story with more detail that draws me into the matter. But he is accountable to others who share the rule, and have well-grounded normative and empirical expectations about Alf's eating habits. In the Open Society, of course, Alf can withdraw from many such moral networks, as others can join them. This is an engine of moral change.

This is a critical point. Philosophers usually see morality as homogenous—and each is accountable to everyone else for failing to act as morality requires. When we look around us, we see the moral world is not at all like this. A normal university professor participates in a wide array of moral networks. At the university as a scientist she participates in networks that have high expectations about evidence and impartiality; a student who flouts them will be the subject of intense moral criticism and, perhaps, severe punish-

ment. But as a member of a church she does not hold others accountable to them: it would be inappropriate for her to start condemning her pastor for lax evidential standards in his Sunday sermon. She may be a vegetarian and hold her university circle of vegetarians to high standards, but she would display outrageous behavior if she walked into a Subway and commenced to berate customers' choices. This is not to deny that the professor might believe that her pastor really ought to pay much more attention to evidence, or that everyone ought to refrain from eating meat. Her *judgments* about morality may or may not be universal. But this is to say that the conditions to hold the pastor or the Subway customers accountable do not obtain, and so their actions are not her business—that is, *they are not accountable to her* for their violations. We navigate these various networks of accountability constantly—so unconsciously that we may be surprised that is what we are doing.

Trendsetter Networks

Although *Tyranny* is critical of the idea that republican communities of inquiry should form “moral communities” (as I have defined them), we must distinguish a moral community from a group of individuals, perhaps who are joined in a republican community of inquiry, who seek to establish a new rule on some matter (Bicchieri 2016, chap. 5). In a polycentric moral order this innovative activity will be restricted to a single moral rule or a small set. Those adopting this new rule will be enmeshed in many other moral rule networks, and so constantly confronting diverse perspectives. They will not form a moral rule community.

On some matters a polycentric order can contain competing moral rules. Given such competition a trendsetter group (say, university students in the '60s), begin experimenting with a new rule (say, concerning sexual morality), withdrawing (usually very publicly) their allegiance to the old. In the case of sexual morality the new networks expanded but did not go to fixation—in many towns and rural areas, and in much of the south of the United States, the old rules of sexual conduct held pretty firm. Moral innovation on some rule of social morality certainly can, then, occur by spreading out from a trendsetting network (that may also be a republican community of inquiry). As Robert Boyd and Peter J. Richerson (2005) have argued, under some conditions group beneficial rules can spread very fast. Sometimes all of society will cascade to a new rule; at other times this competitive process leads to a polycentric order in which different sub-networks adopt different rules (Gaus, 2017c).

Vallier's essay is especially valuable in stressing how moral innovation requires moral space—a protection from universal accountability—for moral diversity and experimentation. So very often today those advocating moral change take their role as requiring that they browbeat others to conform to the moral rules they and their core network are convinced are correct (Gaus 1996, 123ff). They hold the world accountable for not living up to their convictions. An aggressive moral self-righteousness is often seen as mandatory for anyone who seeks a more perfect justice. Anything else is “relativism” or, to again harken back to the '60s, liberal “repressive tolerance.” That, however, is to ossify one's current understanding of justice and to undermine the social conditions for knowing a more perfect justice. We are not at the end of history—and that includes our knowledge of a perfectly just social state.

The Fundamental Rules of the Moral Constitution

Some rules, however, are so fundamental to cooperative social life that it is very difficult to have sustained interactions with those who do not adhere to the same ones we do. If some reject our rules about harm to bodily integrity, property, truthfulness, etc. it will be immensely difficult to share a social life with them. These fundamental rules require a different analysis of moral reform than the competing networks account. When it comes to the rules of sexual morality, university students could practice their own rules freed of the hang ups of the rest of the society (and eventually convert a good deal of it). Sexual relations are certainly social, but even in the '60s they weren't especially large-scale social phenomena.⁶ Some can go their own way. But, leaving aside retreating to a commune, it proved impossible for the '60s trendsetters to informally change property rules. Given our own deep commitments, we have strong reason to share a basic framework of such moral rules, but given the diversity of our moral perspectives we disagree about what those rules should be.

The Order of Public Reason introduced the idea of a socially eligible set of rules, which *Tyranny* employs. The basic idea is straightforward. In a large diverse social network (the Great Society, for example), individuals have clashing views about the specific form these moral rules should take (what should be the rules about harm? what should be the rules about property?). However, these types of rules are so fundamental to an ongoing scheme of moral accountability—which itself is fundamental to effective social life—that almost all individuals are prepared to endorse and adhere to specific versions that fall a long ways short of their most fa-

vored formulation. The socially eligible set, for any specific matter to be governed by a moral rule in a particular social network, is the set of all the rule variations being advanced that everyone (or, as near as possible, everyone, a point to which I shall return) in the network has sufficient reason to conclude are at least minimally worthy of endorsement and adherence. Since for these fundamental rules we really need to coordinate on a single, shared, rule, the socially eligible set identifies the set of possible rule variations all can understand as grounding a shared practice of moral accountability. For every rule in the socially eligible set, each person would endorse and adhere to it should it be the rule that their society (network) has hit upon.

How, then, can an Open Society change such rules? As I argued in *The Order of Public Reason* democratic government reforms the fundamental rules of our society through legislation. When the law moves us within the socially eligible set, the result is still a rule that everyone holds is a basis of genuine moral accountability, but many believe is a better rule (given their moral perspectives) than the one it replaced. Democratic decision making is, I think, indispensable is reforming our truly fundamental rules. In *Tyranny*, however, I explored informal, social mechanisms by which social movements can seek to move us to what proponents see as a superior basis for accountability. One interesting avenue I consider is to exploit rule ambiguity. All rules are ambiguous in many places, and typically in such situations others will accept several alternative actions as plausibly fitting the rule. In these circumstances we can nudge the rule in directions we morally favor (in the eligible set), without denying the validity of the normative expectations of others. For example, when I was growing up the basic rules about physical violence toward children allowed corporeal punishment, but not too much violence. The ambiguity about what constituted unacceptable violence toward children allowed a change in the basic rules, as some began to employ increasingly stringent interpretations of what “violence” was and when it was unwarranted. This was extraordinarily effective in changing the moral rule. When I was a lad seeing a mother smack her child in the supermarket was not especially rare or noteworthy. Now it is clear violation of a basic rule.

The Limits of Moral Space

In an earlier paper D’Agostino (2013) worried that, instead of finding an “eligible set” of acceptable moral rules, the “null hypothesis” will hold—there will not be *any* rule all deem eligible—worthy of endorsement. His paper in this

symposium develops this worry: there seems to be precious little chance of an eligible set in a society split into opposing perspectives that “are increasingly likely to treat those whose adopt different social ideals as pariahs, unworthy of moral regard.” The obvious case here is the great animosity in the United States between so many Republicans and Democrats.

D’Agostino is surely correct that this dismissal of the moral status of others is one of today’s most serious threats to the Open Society. However, while recognizing the danger this poses to the Open Society, I also think we should be aware that to a large extent this a political problem more than a generalized moral one (reflect on the obvious example). As Milton and Rose Friedman (1980, p. 66) long ago pointed out, political decisions require “conformity without unanimity” whereas self-organizing systems (like the market) produce “unanimity without conformity.” In a highly morally diverse society, when political decision-making pushes beyond maintenance of core rights and liberties to the legal codification of deeply controversial conceptions of justice, hostility and contempt for the law is apt to be triggered (Gaus 2017a). Politics is ill-equipped to cope with deep moral disagreement (i.e., where the null hypothesis holds). Each party, hopeful that a majority win in the next election will allow it to institute true justice, simply sets the stage for the next iteration of mistrust. To argue that democratic societies need to develop more trust, while seeing them as a struggle about which controversial conception of justice will be imposed on the appalled minority, is ultimately incoherent. One cannot make politics a justice jihad and hope to induce trust (see Vallier, forthcoming).

The idea of a polycentric moral order helps to show how we can secure “conformity without unanimity.”⁷ Although we are currently understandably obsessed by the hatred underlying so much American politics, we should not forget that Democrats and Republicans share a myriad of rules about bodily integrity, property, gender equality (yes, though they disagree on the policies to pursue it). They cooperate in neighborhood organizations (my own neighborhood, for example, has about an equal distribution of hybrids and pickups, yet an active neighborhood organization). The more our moral rules track networks of individuals seeking to live together and solve their social problems, the less the null hypothesis will be a worry.

I do not want to seem Pollyannaish. The debate about the status of abortion rights is a deep moral disagreement that inevitably flows to political dispute. Even about this vexed issue, however, I think a more decentralized politico-legal

system that is responsive to the differences of locality and region could at least mitigate the depth of hostility (Gaus 2017a; Kogelmann 2017, chap. 4). A morally diverse society is bound to have deeper political conflicts than one that gives the appearance of moral homogeneity (because some perspectives have been silenced or have been dismissed as unreasonable). Politics amplifies disputes, and these can obscure well-functioning informal networks on which an extended order of cooperation is built. Perhaps the main message of my last two books has been the importance of the social and informal, which contemporary political philosophy almost wholly ignores.

Still there are limits. I agree with Rawls in holding that the moral space of endorsed cooperation is always limited. Some will simply refuse to live with others on terms that ground mutual accountability and mutually endorsed expectations. Recall, though, that Rawls (2005, pp. 197-8) credits Isaiah Berlin with this thought. Whereas Neufeld and Watson see it as bound up with the Rawlsian ideas of reasonableness and reciprocity, I see it as simply following from the recognition of Berlin's core theme of deep diversity. Given a deep enough diversity of moral perspectives, it is inevitable that some won't be able to see their way to endorsing the rules of cooperation that almost all others employ to structure their social-moral lives. Neufeld and Watson, as Rawlsians, would tend to describe these folk as "unreasonable." Of all the Rawlsian categories, this is the most vexed. While I do not believe it is hopelessly vague or confused, I certainly do not wish to employ it in any formal way. As I see it, typically these "excluded perspectives" are those who value their own purity (or, perhaps we should say, less controversially, their integrity) over reconciliation and living with others. I have recently modeled them as "maximum integrity agents" (Gaus, 2017c). These may be devoutly religious folk, or Kantian moral philosophers. I am not prepared to say the maximum integrity stance renders one unreasonable, but it does tend to make one unfit to live with many others. One need not embrace reciprocity to find a path to reconciliation with the moral perspectives of others: that is one route, but there are many—yet some may fail to find any of them.

However, we must remember that, *pace* the social contract tale, living with others on moral terms is not a single decision, such that one is "either in or out." That social-moral life is so clearly *not* like that should lead us to question the entire social contract approach. No one except a sort of cartoon nihilist (more plausibly, a psychopath) is a "holdout" on social morality *per se*, but rather one "holds out" on this

or that moral rule in this or that network. Some who insist that the very idea of living with diverse others is an insult to their integrity will find themselves with a shrunken and impoverished moral space, but they will participate in some networks. However, they will not be able to reap the benefits of the Open Society. And the rest of us will have to beware of them and probably guard against them, as they may seek to undermine the basis of our public moral world.

IV. THE IMPERATIVE OF COMPLEXITY AND THE NEW PROGRAM FOR POLITICAL PHILOSOPHY

In 1971 Rawls's *Theory of Justice* revolutionized political philosophy, taking as its subject principles for a society characterized by enduring rational, normative disagreement. Today, in a society in which basic facts are in dispute and perspectives face each other with mutual incomprehension, the idea that all rational moral agents would agree on the difference principle⁸ as regulating social and economic inequalities seems rather quaint. Rational moral agents disagree on the good life, but not about social justice! We can only try to remember when that seemed like deep disagreement. To his great credit, Rawls continually explored the basis of disagreement about the justice of our society, leading him deeper and deeper into the problems of social organization under diversity. Yes, I believe that at the end his political liberalism project was in disarray, in the sense that its unity and organization broke down (I would certainly not say, as Neufeld and Watson think I would, that it was a "jumbled mess"). As Chad Van Schoelandt and I (2017) argue, this disarray is a testimony to the protean nature of Rawls's project. He was constantly inventing new terms and advancing new analyses to capture evolving insights. I see no point in freezing an ongoing project at a moment in time, and insisting that it was finished. It manifestly was not.

What Rawls entirely failed to appreciate, however, was that as we make a system of interaction increasingly diverse, not only does the basis for consensus become increasingly thin, the system becomes increasingly complex. Diversity and complexity are intimately related (Page 2011). To accept that the subject of political philosophy is a system characterized by the dense interaction of diverse moral agents leads to the conclusion that its subject is in the formal sense a complex phenomenon. It is this "imperative of complexity" that is at the heart of the new program in political philosophy that D'Agostino announces in his essay. As Hayek

insisted throughout his career, constructivist blueprints for complex systems often look wonderful on the drawing board but simply cannot be built. The new program accepts the imperative of complexity, and so focuses elsewhere: on the perspectives of diverse moral agents and under what conditions they can organize themselves into a fruitful and cooperative social life that all endorse as moral—given their heterogeneous understandings of that contested concept. As I understand it, this new program switches focus from the moral convictions and plans of the philosopher to those of the agents who form the self-organizing and self-governing moral order. Once they abandon the legislative and planning tasks, political philosophers have much to contribute, especially if they are willing to engage in trans- and inter-disciplinary inquiry. In place of conceptions and plans for justice the new program—or at least my version of it—seeks to identify “devices” of public reason, to analyze how markets, democracy, polycentrism, liberty rules and jurisdictional rights provide the framework for diverse moral perspectives to form moral and political orders that not only accommodate, but improve, their disparate understandings of justice in an open society.

NOTES

- 1 For useful summaries, see Page’s essay and Mitchell 2009, pp. 281-6.
- 2 I write here as a great fan of Betamax, being among the last to abandon it.
- 3 Sometimes simple optimization models seek to provide a non- V based structure among the members of X through appeal to feasibility; it is offered as a way to say what options are “close” and “far” in a sense independent of the value function. *Tyranny* argues that while feasibility is an important and often critical idea, it does not have the characteristics to define coherent and well-behaved structural relations among the members of X . We cannot build a better coherent structure by asking, at each moment, what improvements on the current structure are the most feasible.
- 4 More formally, I think we would want our binary-based theory of the ideal to satisfy some version of expansion and contraction consistency. See Sen 2017, pp. 317-23.
- 5 Thus the idea on which Vallier focuses: that a violation of a moral rule in one’s network is “everyone’s business.” To be sure, I used this as something of a motto rather than a strict necessity, as a network can develop moral rules in areas like marriage (e.g. fidelity), but interests in privacy may lead to the conclusion that violations are only the business of the family unit. Nevertheless, without further considerations the general statement holds within any given moral rule network. In economic terms, rule violation is a public bad, so all are concerned when violations occur. Except for the special case of human rights, I did not, however, hold that conformity is the business of those outside the network (in *The Order of Public Reason*, outside of G). As I stressed, other groups may have different moral rules, and in most cases (human rights aside) one cannot hold them accountable for failing to conform to our social morality. So, yes, within some given G , violations are almost always a concern to all.
- 6 Although according to what my daughter calls “the source of champions,” *Wikipedia*, up to 100,000 people participated in the summer of love.
- 7 Brian Kogelmann (2017) has rigorously analyzed the proposal that the political system draw on polycentrism to secure this.

- 8 That is, social and economic inequalities should be arranged so that they maximally benefit the least well off class.

REFERENCES

- Bicchieri, C. (2016). *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*. New York: Oxford University Press.
- Boyd, R. and Richerson, P.J. (2005). Group-Beneficial Norms Can Spread Rapidly in a Structured Population. In: *The Origin and Evolution of Cultures*. Oxford: Oxford University Press, pp. 227-40.
- Colander, D. and Kupers, R. (2014). *Complexity and the Art of Public Policy: Solving Society's Problems from the Bottom Up*. Princeton: Princeton University Press.
- D'Agostino, F. (2010). *Naturalizing Epistemology: Thomas Kuhn and the 'Essential Tension.'* London: Palgrave-McMillan.
- (2013). The Orders of Public Reason. *Analytic Philosophy*, vol. 54 (March): 129–155.
- Downs, A. (1956). *An Economic Theory of Democracy*. New York: Harper and Row.
- Friedman, M. and Friedman, R. (1980). *Free to Choose*. London: Secker and Warburg.
- Gaus, G. (1996). *Justificatory Liberalism*. Oxford: Oxford University Press.
- (2011). *The Order of Public Reason*. Cambridge: Cambridge University Press.
- (2016). *The Tyranny of the Ideal: Justice in a Diverse Society*. Princeton: Princeton University Press.
- (2017a). The Open Society and Its Friends: With Friends Like These, Who Needs Enemies? *The Critique* <<http://www.thecritique.com/articles/open-society-and-its-friends/>>. Accessed May 23, 2017.
- (2017b). Social Morality and the Primacy of Individual Perspectives. *Review of Austrian Economics*, vol. 30: 377-396.
- (2017c). Self-organizing Moral Systems: Beyond Social Contract Theory. *Politics, Philosophy and Economics*. DOI: 10.1177/1470594X17719425
- (2018). The Priority of Social Morality. In *Morality, Governance, and Social Institutions: Reflections on Russell Hardin*, edited by Thomas Christiano, Ingrid Creppell and Jack Knight. New York: Palgrave McMillan, 2018.
- and Van Schoelandt, C. (2017). Consensus on What? Convergence for What? Four Models of Political Liberalism. *Ethics*.
- Hayek, F. A. (1964 [2014]). The Theory of Complex Phenomena. In: *The Market and Other Orders*, edited by Bruce Caldwell. Chicago: University of Chicago Press, pp. 257-266.
- Hess, M. B. (1966). *Models and Analogies in Science*. Norte Dame: Notre Dame University Press.
- Hotelling, H. (1929). Stability in Competition. *Economic Journal*, vol. 39: 41-57.
- Johnson, James (2014). Models Among the Political Theorists. *American Journal of Political Science*, vol. 58 (July): 547-560.
- Knight, F. (1921 [1964]). *Risk, Uncertainty, and Profit*. New York: Augustus M. Kelly.
- Kogelmann, B. (2017). *Agreement, All the Way Up: An Essay on Public Reason and Theory Choice*. Doctoral dissertation, University of Arizona.
- Kuhn, T. S. (1970). *The Structure of Scientific Revolutions*, second edn. Chicago: Chicago University Press.
- Lane, R. (2017). *The Complexity of Self Government: Politics from the Bottom Up*. New York: Cambridge University Press.
- Levy, D. M. and Peart, S. J. (2017). *Escape from Democracy: The Role of Experts and the Public in Economic Policy*. Cambridge: Cambridge University Press.
- Mitchell, M. (2009). *Complexity: A Guided Tour*. Oxford: Oxford University Press.
- Nozick, R. (1974). *Anarchy, State and Utopia*. New York: Basic Books.
- Ostrom, E. (1986 [2014]). An Agenda for the Study of Institutions. In: *Choice, Rules and Collective Action*. Eds. Filippo Sabetti and Paul Dragos Aligica. Essex: ECPR Press, pp. 97-119.
- Ostrom, V. (1972 [1999]). Polycentricity (Part 2). In: *Polycentricity and Local Public Economies*. Ed. Michael D. McGinnis. Ann Arbor: The University of Michigan Press, pp. 119-38.
- Ostrom, V. and Ostrom, E. (1977 [1999]). Public Goods and Public Choices. In: *Polycentricity and Local Public Economies*. Ed. Michael D. McGinnis. Ann Arbor: The University of Michigan Press, pp. 75-103.
- Page, S. E. (2011). *Diversity and Complexity*. Princeton: Princeton University Press.
- (2016). Not Half Bad: A Modest Criterion for Inclusion. In: *Complexity and Evolution: Toward a New Synthesis for Economics*. Eds. David S. Wilson and Alan Kirman. Cambridge, MA: MIT Press, pp. 319-26.
- Poundstone, W. (1992). *Prisoner's Dilemma: John von Neumann, Game Theory, and the Puzzle of the Bomb*. New York: Random House.
- Rawls, J. (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- (2005). *Political Liberalism*. Expanded edn. New York: Columbia University Press.
- (2007) *Lectures on the History of Political Philosophy*. Ed. Samuel Freeman. Cambridge, MA: Harvard University Press.
- Schumpeter, J. A. (1950). *Capitalism Democracy and Socialism*. Third edn. London: George Allen and Unwin.
- Sen, A. (2017). *Collective Choice and Social Welfare*. Expanded edn. Harmondsworth: Penguin.
- Vallier, K. (forthcoming). *Must Politics Be War? In Defense of Public Reason Liberalism*. Oxford: Oxford University Press.
- Weisberg, M. (2013). *Simulation and Similarity: Using Models to Understand the World*. Oxford: Oxford University Press.