BETWEEN DISCOVERY AND CHOICE:
THE GENERAL WILL IN A DIVERSE SOCIETY

GERALD GAUS
ggaus@email.arizona.edu
University of Arizona

ABSTRACT. In his *Lectures on the History of Political Philosophy* Rawls tells us that "a normalization of interests attributed to the parties" is "common to social contract doctrines." In the first part of this paper I argue that this "normalization contract" faces a dilemma. If the normalization theorist insists on strict normalization, he can get agreement about how to promote the common good, but we will almost certainly be stuck at poor articulations of it; once the normalization theorist allows diversity, he is unable to contain it to "just the right amount" – the very diversity that he needs to escape being caught in poor local optima undermines the strong normalization thesis. In the second part of this paper I sketch two mechanisms by which free and equal persons can come to endorse a general will even when they do not agree what is best: political procedures and social equilibrium.

Keywords: general will, perspectives, diversity, social contract, public reason, common good

1. The Dilemma Facing the Normalized Social Contract

In his *Lectures on the History of Political Philosophy* Rawls tells us that "a normalization of interests attributed to the parties" is "common to social contract doctrines."[1] This remark is made in the context of discussing Rousseau's idea of the general will, which is also said to require a shared "point of view."[2] On Rawls's reading of Rousseau, as private individuals we have a variety of different interests and are characterized by self-bias and selfishness. Such individuals can live together under freely endorsed common laws only if they "share a conception of the common good."[3] Such a shared conception is generated by their fundamental interests and capacities, which derive from their shared human nature. As Rawls sees it, these shared fundamental interests allow the parties to abstract from their

differences and occupy a shared legislative point of view, based on a shared conception of the common good.[4] When occupying this shared view they all have the same basis for their deliberations, and so will the same laws: and that is why they freely legislate common laws.

The conviction underlying this reading of the social contract tradition and the general will is that a diversity of perspectives threatens the project of uncovering agreed-upon moral and political principles. If we are to have any prospect of arriving at a general will, we must abstract away from our differences, reasoning only from a shared, "normalized," point of view. In the first part of this paper I argue that this "normalization contract" faces a dilemma. If the normalization theorist insists on strict normalization, he can get agreement about how to promote the common good, but we will almost certainly be stuck at poor articulations of it. In the more technical term that I shall employ, we will be stuck at "poor local optima." I shall show that if we are to uncover a political and social order that can be reasonably expected to advance all the constituent elements of the common good, we must allow a diversity of perspectives into deliberation about the common good. However, once the normalization theorist grants this, he is unable to contain diversity to "just the right amount" – the very diversity that he needs to escape being caught in poor local optima undermines the strong normalization thesis. In the second part of this paper I sketch two mechanisms by which free and equal persons can come to endorse a general will even when they do not agree what is best: political procedures and social equilibrium. Thus the first part of the paper concerns how we maximize our chances of reaching an optimal general will; part two focuses on what to do when we confront what Amartya Sen has called a maximal set with no optimal element.

A plausible social contract must be a diverse, not a normalized, contract.[5] A normalized contract, I hope to show, is unpersuasive in its own terms and, more importantly, the social contract so conceived is unsuitable as the basis of a diverse society of free and equal persons. It only sees diversity and difference as problems to be coped with (and that by refusing to seriously think about them), not resources to be employed in justifying our laws and social rules. Once we understand the role of diversity in reasoning about the common good, we shall see that the general will in a diverse society is a matter of both discovery ("Just what norms and laws promote the common good?") and a matter of social choice ("Given that we have different views of this question, how can all free and equal persons come to endorse common norms and laws?"). Such a contract of "bounded diversity," gives us genuine insight into how a diverse society can come to share a general will, and so a common moral life.

Part One. Normalization

2. Why Normalization Is the Problem, Not the Solution

2.1 The Elements of a Normalized Perspective

The entire normalization view of the social contract – and, more generally the contemporary Rawlsian public reason project – supposes that *if* we could only achieve normalization, *if only* the problem of social evaluation could be reduced to reasoning based on a single perspective, then we would have solved the problem of uncovering the social rules (laws, institutions, or whatever) that would promote the common good. To be sure, there may be costs incurred and controversial moves made along the way (i.e., the original position with its veil of ignorance), but these would be well worth it if they could identify a shared perspective for evaluation. Before asking whether this normalization is plausible, let us ask: if we could indeed achieve it, would it be desirable?

To "share a point of view" or a "perspective" is not simply to (*i*) share a set of fundamental interests or values, but involves also sharing (*ii*) a common way to map these values on to social worlds (sets of social arrangements, basic structures, etc.) and (*iii*) a shared understanding of the social world subject to evaluation.[6] The idea of evaluating a "social world" may initially strike us an unfamiliar or appear to imply that we directly choose over social outcomes.[7] However, the idea is neither alien nor narrow in its assumptions; it supposes simply that whatever choices we make about our overall set of institutions, rules, or practices results in an overall social state that we wish to evaluate.[8] Consider in this light Rawls's exposition of Kant's first formulation of the categorical imperative (the universal law formulation) in terms of the four-step CI procedure:

(1) I am to do *X* in circumstances *C* in order to bring about *Y*. (Here *X* is an action and *Y* a state of affairs).

The second step generalizes the maxim at the first to get:

(2) Everyone is to do *X* in circumstances *C* in order to bring about *Y*.

At a third step we are to transform the general precept at (2) into a law of nature to obtain:

(3) Everyone always does *X* in circumstances *C* in order to bring about *Y* (as if by a law of nature).[9]

In this last step we are to consider the "perturbed social world" that would result from the addition of this new law of nature; we seek to understand the new "equilibrium state" on which this perturbed social world would settle. We are then to ask ourselves whether, when we regard ourselves as a member of this new social world, we can "will this perturbed social world itself and affirm it should we belong to it."[10]

If, then, we secure agreement on (*i*)-(*iii*), contractor-deliberators share a fully normalized perspective and can proceed to the core question of social contract theory. Given (*i*) and (*ii*) what set of laws and institutions governing (*iii*) would best promote the common good? Points (*ii*) and (*iii*) are essential if the shared perspective is to succeed in its aim of producing a shared general will via agreement in judgment. Consider (*ii*): it would be of no avail to share a set of interests or values if deliberators employed different aggregation methods or metrics that determine an overall score or judgment for some combination of values. Suppose the normalized contractors agree that V_1, V_2 and V_3 are the relevant values included in the common good that relate to the evaluation of social worlds *x*, *y* and *z*. Even suppose that in this case that they perfectly agree how each social world scores on each of these, as in Display 1.

	x	*y*	*z*
V_1	Low	Medium	High
V_2	Medium	Medium	Low
V_3	High	Low	Low

DISPLAY 1 AN EXAMPLE OF UNAGGREGATED EVALUATIONS

If our contractors are to agree in their overall evaluations, they need to use a common method to aggregate their judgments into an evaluation (e.g., an ordering) of the various social states under consideration. Call this the *mapping function*. The mapping function would take a set of evaluative standards $\{V_1, V_2 ... V_n\}$ and produce an ordering of a set of alternatives – social worlds.

2.2 Social Worlds in our Neighborhood: The Feasible Set

It is essential to realize that for the normalization assumption to perform its task of producing agreement, the parties must share the same understanding of what constitutes the social worlds to be evaluated, which means consensus on what parts of it are relevant to the common good (our third requirement). A social world, we might say, is composed of both a sort of "normative ontology" and certain understanding of social and psychological facts that must be taken as given in social evaluation. The normative

ontology concerns such questions as whether the common good is to be applied to the public and not the private; and if so, what parts of the social order are public and which are private?[11] Is it to be applied to the basic structure – and what does the basic structure encompass? Is the family part of it? Is the common good to regulate the political, the social, or the personal too? A perspective thus identifies the aspects of the social world that are relevant to evaluation. It also identifies the background facts that must be taken as given in social evaluation. One of the features of the famous dispute between Rawls and G. A. Cohen – about whether parties to the original position should suppose that the principles of justice must provide non-moral incentives for contribution to the common good – concerns a difference in the social worlds that are subject to evaluation.[12] Rawls's social worlds seem to many of us more proximate to our own, where these psychological facts are taken as given, whereas in Cohen's the parties are to evaluate the social worlds in which compliance via moral incentives is sufficient to ensure that people will conform to the dictates of distributive justice. A social world sums up all that, for the purposes of normative evaluation, is to be taken as in some way fixed, and determines what is to be evaluated.

In *Justice as Fairness: A Restatement* Rawls advances a conception of alternative social worlds in a "neighborhood" in the process of responding to an objection of Derek Parfit's to the difference principle, as illustrated in the example in Display 2.[13]

	Indians	British
(1)	100	100
(2)	120	110
(3)	115	140

DISPLAY 2 A PROPOSED COUNTEREXAMPLE
TO THE RECIPROCITY CLAIM OF THE DIFFERNCE PRINCIPLE

The difference principle selects distribution (3) because the least well off do best. But Rawls claims that a justification of the difference principle is that the shares of the better off are not gained at the expense of the least well off. As Rawls stresses in his later work, the difference principle expresses reciprocity, a commitment of the better off not to gain at the expense of those who are already less well off. Yet we see that under distribution (2) the Indians do better than they do under (3), so it would see that the gains for British under distribution (3) come at the expense of the Indians, who "lose" 5 units.

Rawls's reply is multifaceted; he insists that the difference principle does not refer to rigid designators such as "Indians" and "British" but to whomever is the least well off class. However, he continues on:

> Ignoring the matter of names for a moment, consider what can be said to the Indians in favor of (3). Accepting the conditions of the example, we cannot say that the Indians would do no better under any alternative arrangement. Rather, we say that, in the neighborhood of (3), there is no alternative arrangement that by making the British worse off could make the Indians better off. The inequality in (3) is justified because in that neighborhood the advantages to the British do contribute to the advantages of the Indians. The conditions of the Indians' being as well off as they are (in that neighborhood) is that the British are better off.
>
> This reply depends, as does the difference principle itself, on their being a rough continuum of basic structures, each very close (practically speaking) to some others in the aspects along which these structures are varied as available systems of social cooperation. (Those close to one another are in the same neighborhood). The main question is not (3) against (2), but (3) against (1). If the Indians ask why there are inequalities at all, the reply focuses on (3) in relation to reasonably close and available alternatives in the neighborhood. It is in this neighborhood that reciprocity is thought to hold.[14]

Rawls is here concerned with a continuum of basic structures – social worlds – in terms of distributional possibilities. But we may, and I think, should, press the idea of such a continuum further, in terms of neighborhoods of social worlds that, we might say, differ in their social-normative ontologies. For example, in Rawls's normalized contract the family sits (somewhat uncomfortably) in the basic structure.[15] The question arises: can our normalized Rawlsian contractors even *think* about whether it would be better, to say, exclude the family from the basic structure? Habermas thought not; on his reading, Rawls relied on an "a priori" conception of the basic structure (for example, concerning what is public and what is private) that was exogenous to liberal justification.[16] Rawls disagreed: A "sphere of life" is not "something already given apart from political conceptions of justice."[17] So it seems clear that the deliberators should consider the benefits of conceiving the social world in different ways, at least those that are in the same neighborhood. They could come to wonder whether private organizations such as churches should be included in the basic structure, and so subject to the requirements of fair equality of opportunity.

2.3 Does a Fully Normalized Perspective Imply a Metric of Distance? (Yes)

Allowing that our fully normalized deliberators can think about different social possibilities (different types of social worlds in which, say, the objects for evaluation are different or people's motivations are different), Rawls is quite clear that they conceive of these in terms of a metric of distance, thus the concept of basic structures within a neighborhood. Those that are near to each other are in the same neighborhood, whereas those that are further (say, where the Indians are the better off class and the British the worse off) are quite distant. As Rawls's reply to Parfit shows, from the perspective of justice as fairness we suppose some set of feasible alternatives that are in our neighborhood, and justify a social world as the best in the neighborhood. The obvious way to think about this is that contractor-deliberators accept that, say, S_X is their social world (where Indians are the less well off) which is to be subject to evaluation, and consider movements from S_X in their neighborhood. Now we confront a core issue: suppose our fully normalized reasoners consider two possible alternative social worlds to be evaluated, S_Y and S_Z. We must consider three possibilities:

(*i*) they share a view of how great a distance S_Y and S_Z are from S_X, i.e., they agree either that S_Y is closer or further than S_Z or that they are equally close;

(*ii*) they do not share a view about distance in the sense that some hold S_Y is closer to S_X than S_Z, and others hold that S_Z is closer, etc.;

(*iii*) they make no judgments of distance at all.

Rawls's reply to Parfit is, essentially, a rejection of (*ii*) and (*iii*); it is because we all recognize that distribution (2) is so far from the present social world that we can eliminate it as irrelevant. Rawls thus endorses (*i*). If the normalized contract is to help us to decide in a uniform manner about relative evaluation of social words in a way that takes account of the distinction between those that are in the same neighborhood and those that are not, we must suppose (*i*), a common ranking of social world distance.

So to share a "point of view" is to share a perspective in the sense that, in relation to our world S_x, some alternative social worlds appear quite close (distribution [1] is close to [3]) and others (distribution [2]) distant. Imagine, then, that our contractor-deliberators who share a perspective in this fully normalized sense consider different social worlds: basic structures that include less or more, which change the objects of evaluation, or our basic assumptions about society. From their fully normalized perspective they can ask "would the common good fare better if we changed our conception of the basic structure?" Because they are asking this question

from their fully normalized perspective, a common understanding of *social world distance* follows, as Rawls suggests, from full normalization. If the contractors share all three of the elements of a common perspective (a common set of values, a mapping function, and a common understanding of what we now might call "their present social world") they will agree on how to arrange worlds in terms of proximity. Agreeing on all these matters, suppose that we are at S_x right now; we will also agree on what is very close to our present arrangement and what is very far.

2.4 The Deliberators' Optimization Problem

So let us say that our deliberators can consider a range of alternative social worlds $S_1...S_n$, and ask which they would prefer to others. One way of thinking about this is for them to compare the principles of justice, or common-good promoting set of institutions, that they would choose for each social world (thus they give each social world a common good score). We now arrive at the crucial problem: sharing a fully normalized perspective almost surely means deliberators will fail in identifying a social world that best satisfies the requirements of common good – what we might think of as a conception of the best feasible social world (we call it "a realistic utopia"). Our deliberators confront an optimization problem: given the normalized perspective that relates different dimensions of the common good (the different fundamental interests), and maps them on to social worlds $S_1...S_n$, what social world optimizes over all the dimensions of the common good?

Suppose we have N dimensions of the common good that display K interdependencies. If optimizing the common good is this sort of complex problem, we are confronted with an NK optimization problem: in evaluating improvements we are optimizing over N dimensions with K interdependencies among them. For example, the common good value of a certain amount of equality (e) may vary widely with the freedom we have: e with no freedom may set back the common good, while e with high freedom (f) may be highly desirable, such that $\{e,f\}$ greatly exceeds the additive values of e and f alone. Still, there sometimes may be too much freedom and equality, such that combinations of extreme amounts of both are harmful to the common good.[18] If the dimensions of the common good are multiple and interdependent in this way then the evaluators are faced with the sort of complex decision problem recently analyzed by Scott Page, Fred D'Agostino and others.[19]

When $K=0$, that is when there are no interdependencies between the dimensions, local optimization decisions (optimizing within our own neighborhood) will put us on a path to global optimization or ideal justice.[20]

However, as the value of *K* increases a rugged optimization landscape emerges. To see the problem, contrast Displays 3 and 4.

In Display 3, the extent to which a social world satisfies value α is mapped on the vertical axis. The horizontal axis employs a perspective, which ranges social worlds in terms of their proximity. (How alike are these social worlds, given this point of view?) For example, suppose the point of view is the traditional left-right perspective, in which social state S_1 is the far left and social state S_n is the far right, and suppose α is the value of individual freedom. Suppose that we are now at X. Note that in Display 3, for any social state (except at the peak, the global optimum) there is a proximate social world that does better at satisfying α. In this case, the value promotion problem is easily solved: we only have to have knowledge of the social world near our present state of X, knowing this we know what direction to travel – what changes are recommended by α. And at each point along the path, we will be faced with the same happy problem: mere knowledge of the proximate social states is enough for us to make a decision, and the series of small reforms will lead to the global optimum. If a society begins locally optimizing on a Mount Fuji landscape, it ends up at the global optimum through a series of steps, each of which leads to a better social state.
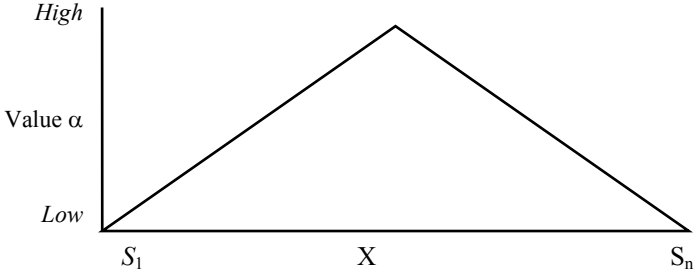
DISPLAY 3 A MOUNT FUJI LANDSCAPE

High

Total
Value
Set

Low

S₁   a   b   u   Sₙ
              X
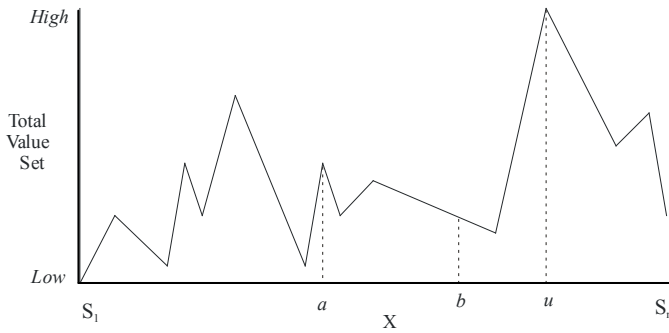
DISPLAY 4 A RUGGED LANDSCAPE

The problem is much more difficult in Display 4. Here the vertical axis maps the total satisfaction of a complex bundle of interrelated values (their relation is not, say, simply additive).[21] Again, the horizontal axis arranges social arrangements according to some perspective. Here, at social world X we are at a local optimum; moving in either direction from X will mean a decrease in overall satisfaction of some our values $\alpha\dots\omega$. This might occur, for example, if at our current world of X value β is highly satisfied, but in ways that preclude the satisfaction of many others (perhaps we have an anarchistic sort of liberty); given the perspective employed, in the near social worlds there would be less satisfaction of, say, value β, but there would not yet be significant gain in other values, resulting in a net loss of total value satisfaction in social worlds adjacent to X.

If under our highly normalized contract where we all share the landscape of Display 4, if we are now in social world X, and if we are boundedly rational, and so only have firm ideas of the evaluation of social worlds in our present neighborhood, we may well conclude that our present social world is optimal: *moving in any direction in our neighborhood will satisfy the requirements of the common good less well than in our present social world*. We are then caught at a local optima, which is considerably short of the global optima. Suppose instead that we are less boundedly rational, having the ability to adequately evaluate a larger neighborhood, say between *a* and *b*. We engage in incremental reform within the bounds of our reason. We will then decide that the common good would fare better at state *a*; should we seek to reform in that direction we would move further from the global optimum! We thus see how real individuals who really share a common (normalized) perspective and have identical epistemic traits are apt to land at a social contract in a social world stuck at a local optimum.

18

3. Minimal Diversity: Allowing Different Search Strategies within the Same Perspective

A society seeking to further the common good – that is, to best articulate the general will – can do better if it takes seriously Rousseau's insight that discovery of the general will is a social rather than an individual (i.e., radically normalized) project, even when we do all share the same perspective. Let us continue to suppose we all share the same perspective, and so occupy the same rugged landscape in Display 4: all agree on the way the core values should be mapped on to social worlds, and we concur on which social arrangements are proximate to our present one, and which are very different (distant). But now suppose that the individuals have different search strategies: they use their boundedly rational capacities to explore different parts of the space of possible social arrangements. Some people better appreciate different social possibilities. For example, some are more inclined to study the current social state, others are conformists and study what most others do, some are risk takers and devote great energy to exploring far-flung and less well-known options, while others are simply utopians who search and search for the global optimum.

Call this "search diversity." Notice that this extent of diversity could in principle exist even under a normalization contract, in which we share the same perspective (the same way of mapping our common values relating to the common good on to a shared viewpoint of social worlds). This solution requires the strict exclusion of search-related values from the normalized perspective: the normalized perspective must not admit of any differences in values, so whatever motivates different search strategies must be outside the perspective. We must, that is, share precisely the same view as to what would count as the best solution, and have exactly the same conception of the feasible alternatives and how they relate to each other in each social world, but devote ourselves to the exploration of different worlds. We certainly can imagine contexts in which this can occur. Within a corporation or hierarchical research team, a central controller can define the relevant shared perspective, and then instruct different teams to explore different parts of the rugged landscape. Even in less centralized contexts there may be a relatively well-defined practice that serves the same function of setting well-defined parameters as to what would constitute the best, second-best, etc. solutions (all share the same view of the landscape); participants explore, say, various solutions to a certain technological problem, some incremental departures from the current standard, others quite radical.

Two problems, however, confront this minimal diversity proposal as a way to avoid poor local optima in advancing the common good. (*i*) As long as we are boundedly rational, when the utopian reports that she has witnessed a Mount Everest general will in some far-flung part of the terrain,

it will be very hard for most of us not to be skeptical. She is reporting about the nature of social arrangements that is a great distance from ours – and we all have good reasons to be suspicious of recommendations to take a Great Leap Forward. The utopian may show us that if we only suppose that people will act out of solidarity rather then self-interest, a fully egalitarian world maximizing our set of values will result, but we may nevertheless not see any way to get there from here. We can only know that proximate worlds may be feasible (this, I take it, is the crux of Rawls's reply to Parfit).

(*ii*) It seems doubtful, moreover, that in political (and, we shall see, more generally social-moral) life the variation in the grounds for the search strategies can be plausibly insulated from the perspective by which we evaluate the results of our searches. People tend to explore different possibilities because they see the options differently. Utopians and conservatives, for example, do not simply have different "search strategies," they view the options differently. For the conservative the utopian proposals are hopelessly far away, while the utopian – or at least the "realistic utopian" – typically sees these as achievable. And that, perhaps, is at least part of the reason why reports by the utopian of the high peaks she has found at different areas of the optimization landscape are apt to be dismissed by the conservative, even when they agree on the set of values to be promoted. At least in political life, our interest in whether the best articulation of the general will is to be found near or far is crucially dependent on our sense of what is near or far – that is, we differ on our view of the relevant social worlds and their proximity. We diverge, that is, in at least one aspect of our perspectives (the third part). (Think again of the debate between Rawls and Cohen on the necessity to consider economic incentives in a theory of justice.)

4. How a Diversity of Perspectives Can Yield a General Will

It turns out that, given our problem of optimizing the general will, this disagreement in perspective is a great thing. If we all had a God's eye perspective, and so could scan the entire landscape, we might know what to do (or at least what to strive for), but boundedly rational individuals cannot survey all possibilities, and they cannot fully credit those who make confident reports about them, or give instructions about how they can be reached. If the global optimum is "far away" – and there are many deep crevices in between where we are and our "realistic utopia" – we will think it better to survey our neighboring peaks and rest content that we need not cross the Grand Canyon *en route* to Mount Everest. That is the attraction of Mount Fuji landscapes: we can climb to the top through a series of steps,

each one of which is better than the one before (so there is no need to drop into a crevasse in order to eventually get higher).

We now come to the crucial point. *Whether we are climbing a Mount Fuji or a rugged landscape is not a brute feature of the social world but a feature of our perspective.* Our perspective is what determines whether social world S_n is near or far from X. On a left-right spectrum they may be distant; but need we see the world through the left-right spectrum? How *should* we see it?

According to what Scott Page calls the Savant Existence Theorem: "For any problem there exist many perspectives that create Mount Fuji landscapes."[22] There always is some arrangement of the options that creates a Mount Fuji landscape; if we find one *that makes sense to us*, what was distant can become near, and our problem of how to optimize a complex common good is transformed from the very difficult into the tractable. Reconceptualizing the problem via a new perspective can make the difficult problem of rugged landscapes into the much easier Mount Fuji problem. This is a striking and important result. Perhaps liberalism itself was such a reconceptualization. At one point western societies faced the problem of which false religions to tolerate. Even in Locke's "Letter on Toleration" he still was struggling with this view: while he thought it would promote the good of the commonwealth to tolerate Protestants, extending toleration to Catholics decreased the common good (an England that tolerated Catholics was very far from his own), as did extending toleration to atheists (perhaps an even further social world).[23] Locke was pushing towards a Mount Fuji liberal landscape in which each additional right of conscience and speech advanced the common good, but there still were ravines. Eventually – with much help from Locke – the early modern problem of which false creeds to tolerate was transformed into the problem of securing freedom of thought and belief. The options were arrayed in something much closer to a Mount Fuji landscape.

The idea is that if we give up on our third feature of a fully normalized perspective (§2.1), and allow individuals to apply the shared values and mapping relations to different understandings of the social worlds under evaluation, we often dramatically increase the possibility of a crucial *discovery*: a new way of understanding the problem of the common good that we *now all see* converts our optimization problem into one with strong Mount Fuji characteristics. And to do that is to open the way to a shared understanding of the smooth upward path from here to utopia. In contemporary parlance, non-ideal theory leads us to ideal theory.[24]

5. Diversity Contagion

This line of reasoning has the great benefit that it analyzes the diversity of perspectives not as a problem to be accommodated but as a resource to be exploited. And this is quite true. A rigid sharing of the same perspective almost ensures that we will be caught at a local optimum that falls far short of the global optimum. Yet, we ought not get carried away with the ability of diversity to produce consensus. For here is its great weakness: the analysis thus far supposes that we can relax the third aspect of normalization (shared conceptions of the social world) and yet keep the value set and the mapping relation fully normalized.

Just as we had to inquire with search diversity where the diversity came from (and we saw that differences in perspectives seem crucial), so too must we press of perspective diversity: why do people come to see social worlds in different ways? Why, for instance, do some insist that the family is part of the basic structure while others insist that is it quintessentially private – and the private is not political? Or, why do some think that the problem of incentives should not get in the way of thinking about justice, while others insist that that *is* the problem for justice? In a wide range of cases, it is the values and empirical beliefs forming the core of their evaluative outlook which leads to diverse perspectives: a liberty-favoring set of values leads one to see the social world in some ways, an egalitarian set of values in others. Political debates about the place of the family or the market are often so intractable because they are not simply based on different views as to how a core set of values applies to different understandings of the social world, but because the different understandings of the social world are inexorably linked to a different set of fundamental values, or at least deep differences about the proper mapping relation from a generally shared set of the evaluation of social orders. In political life value disagreement inevitably occurs even when our perspectives are partly normalized, in the sense that even if we share a list of fundamental interests, we disagree about the way this list is to be mapped on to the evaluation of social worlds.[25]

Part Two. Diversity

6. Bounded Diversity and the General Will

Once we admit this even deeper diversity, while a process of social discovery in a diverse society can help rearrange a rugged landscape into a Mount Fuji-ish one, its deliberations almost surely will show that its members disagree about the value topography. If our fundamental set of

values and mapping relations differ – if we relax the normalization of all three elements of a perspective – what is a peak to you may be a foothill to me: your peak scores high on your mapping function, but is not impressive on mine. Once we allow this much deeper diversity, the idea of a "social discovery of the best solution" fades, for we disagree about what the criteria for that are.

Just as we did not wish to go overboard in rejoicing over the way that a diversity of perspectives can yield discoveries about the smooth path to the global optimum, neither should we go overboard in our pessimism that there is such a thing as social discovery about bad and better social arrangements based on a partly normalized – or let us, "boundedly diverse" – perspectives. Even if we continue to disagree on the topographical details, we may still discover that, for a significant set of shared values in certain contexts (say, concerning the environment, educational reform), the way forward (toward better solutions) is along a certain set of paths. Or at least we may be fairly confident that all our ideas of a global optimum do not lie in some directions.

Given the way diversity of values spreads once we abandon full normalization (which, it will be recalled, locks us into poor local optima) the best boundedly rational and good-willed persons can expect, on the basis of their collective reasoning about a partially shared understanding of the common good, is a partial agreement on what is best. We reason on the basis of diverse perspectives: we have different values, different ways to map the values on to the social world, and we disagree about just what aspects of the social worlds are to be evaluated, as well as which background facts we should suppose as given. To be sure, our views will be dynamic: as the process of social discovery proceeds we may sometimes find ourselves converging on Mount Fuji-like landscapes where it is clear that the next optimizing move is also one that will take us to the best social world. But we should expect that the norm will be, as it were, an incomplete convergence on the nature of the terrain, leaving us with a number of proposals in a set of generally acceptable social worlds, with deep disagreement about which towers over the others. As far as the general will is concerned, our shared social topographic map will have some useful elevation contours and distance metrics, but there will be large gaps (as in old maps where "unexplored territory" sometimes covered a vast area of indeterminate size and shape).[26]

So here is our problem. If we insist on rigid normalization so that all *really* share the same perspective, we are almost certain to get stuck at poor local optima. The general will thus would endorse a social world in which the common good is optimally furthered in its neighborhood but there may be other attainable worlds in which we do much better. If we allow a diversity of perspectives, we can make progress in climbing to the global

optimum. But at the same time, it seems that by admitting a diversity of perspectives as an engine of discovery, we have precluded any hope that boundedly rational people converge on a single shared general will.

If this is right, the best a Rousseauean-inspired conception of the general will can hope for is one that arises out of a process of social interaction and discovery, which helps us avoid being stuck in clearly poor local optima, but which must forgo all hope of arriving at a unique general will. We will be left with a set of possibilities, which will be evaluated differently by different normative perspectives. This idea of a non-normalized conception of the general will thus must give an account of how, given this basic indeterminacy in the process of social discovery, a society can still come to have a general will.

7. The General Will and Political Procedures

7.1 Choice under Incompleteness

Suppose, then, that our problem is Rousseau's: citizens seeking laws that reflect the common good, and we are debating about some issue. Because we do not necessarily share the same conception of the social world, we may not see this issue in the same way (I see it as a tax measure, you see it as about family policy). But suppose we share enough to see it as the same law, and we each have our idea of possible alternatives. Let us further suppose that our deliberators have considered a set of alternative laws. Because we have allowed diversity of perspectives, we can assume that a wide range of options and assumptions will be canvassed. This diversity, as we have seen, will promote social discovery, allowing citizens to identify social worlds and arrangements that do badly in promoting their core values. Given this, if social discovery reveals that proposed law L_X Pareto dominates L_Y – if in all perspectives L_X is a higher peak in the value satisfaction landscape – L_Y will be eliminated from consideration. Our diversity of perspectives will thus yield what Sen calls a maximal set[27] of laws, $\{L_i...L_k\}$ with no optimal element: based on the perspective of each person, all citizen-deliberators agree that every member of the set is preferred to proposed laws outside the set, but because of the differences in their perspective they cannot agree upon a ranking within the set. To be sure, if deliberators were able to agree on some uniquely rational bargaining solution, all would share the same bargain-based ranking the choice worthiness of the options; but given that each citizen does not even know the perspectives of others, this is rather a lot to expect. More fundamentally, there is good reason to conclude that bargaining theory is itself inherently indeterminate.[28] So we shall suppose simply that each evaluates the set on

the basis simply of her own perspective. More formally, let us say each has an ordinal ranking of $\{L_i...L_k\}$. Call O the maximal set or, as I shall call it (departing from Sen's use), the *socially optimal set*: L_i is a member of O if and only if there is no law that Pareto dominates it (i.e. it is not the case that there is some law that all deliberators rank above L_i). Let us also add for sake of exposition that there is also some proposed law that L_i Pareto dominates. This concept of Pareto dominance of course, derives from economics and, more specifically social choice theory. But it is merely a formalization of the key contractual idea of unanimous choice defined in terms of unanimous pairwise choice; we are not importing an alien idea into the social contract but simply formalizing its regulative ideal.

We thus can agree that we wish to select from O, the optimal social set, but from the social perspective (the perspective of the social contract), there is no way to order the members of O. For any two members of O $\{L_i, L_k\}$ the deliberators as a group cannot say whether $L_i \succ$ [i.e., is preferred to] L_k, $L_k \succ L_i$, or whether they are indifferent between L_i and L_k, that is, they have not determined that they are equally good. As Sen points out, they seem to face the same problem as "Buridan's ass": the donkey who was precisely midway between two haystacks and could not decide whether to turn right (x) to eat from one or left (y) to eat from others, and ended up dying of starvation (z).

The less interesting, but more common, interpretation is that the ass was indifferent between the two haystacks, and could not find any reason to choose one haystack over the other. But since there is no possibility of a loss from choosing either haystack in the case of indifference, there is no deep dilemma here either from the point of view of maximization or that of optimization. The second – more interesting – interpretation is that the ass could not rank the two haystacks and had an incomplete preference over this pair. It did not, therefore, have any optimal alternative, but both x and y were maximal – neither known to be worse than any of the other alternatives. In fact, since each was also decidedly better for the donkey than its dying of starvation z, the case for a maximal choice is strong.[29]

Given this assumption of an incomplete ordering, standard decision theory cannot show that there is a rational choice to be made from the socially optimal set; on the standard interpretation our deliberators simply cannot choose among the options because they cannot rank them.[30] But, like Buridan's ass, where z is $\{\neg x, \neg y\}$ our deliberators have reason to prefer any member of O to ending up outside of O, and the status quo may itself be outside of O. Buridan's ass's predicament is the same as that of a deliberator who has arrived at the decision that his reasons point to endorsing either L_i or L_k, but not both, over some third alternative, L_z. At this stage of his deliberation he is stuck at the conclusion that he does not

have sufficient reason to judge that L_i is at least as good as L_k, and he does not have sufficient reason to judge that L_k is at least as good as L_i, though he does have sufficient reasons to conclude that both are better than L_z. The important thing is not to end up with L_z. As Sen remarks, only an ass wouldn't choose L_i or L_k over L_z.

7.2 The Fair and Efficient State

Political procedures are the fundamental device proposed by the social contact tradition for avoiding making political asses of us. Hobbes, Locke, and Kant all trace the basic problem of social life to the conflicting judgments of private reason about morality, and all see the solution to be establishing the state as the final umpire as to the demands of justice.[31] This is not to say that citizens believe that the political procedure yields the correct social ordering of the optimal set, for *ex hypothesi* there simply is no such ordering. But by providing a single authoritative choice it allows the citizens to act in the face of disagreement and, so avoiding the asinine outcome.[32]

In one of its very many senses, we might say that a *legitimate* state is one that chooses from the optimal set through a procedure that all citizens endorse. But because this term is used to describe too many diverse requirements, let us instead refer to a "fair and efficient state": it is fair insofar as its selection procedures are endorsed by all, and it is efficient in promoting the common good insofar as its laws and policies (*i*) select from the socially optimal set and (*ii*) promote processes of discovery such that we come to better identify what options are and are not in the set. In such a state citizens generally will disagree with the outcome of the political process because in their own perspective the outcome is not ranked highest in set *O*, but they also recognize that from the collective perspective of all citizens, none of the members of *O* can be socially ranked, so some citizens will always and necessarily find that the political procedure does not yield the best outcome as defined by their perspective. In this case there can be no grounds for complaint that the procedure has gone awry, even if one thinks the outcome is nowhere near optimal. But so long as the state chooses from *O*, and so long as our political system reasonably exploits opportunities for social discovery by drawing on a variety of perspectives and so refining our understanding of the optimal set, it does as good at promoting the common good as citizens, respecting each other as free and equal, can require.[33]

David Estlund has made famous the idea that democracy may be justified in terms of "epistemic proceduralism" – its procedures are fair and tend to give the right answer.[34] Leaving aside the details of his own account, we can see that an acceptable political procedure must indeed posses both fairness and good result components. Fairness is important because from the

social perspective within the optimal set there simply is no such thing as a "better" or "worse" result. Consequently, there is no question that someone might show that a nondemocratic procedure does a better job at selecting from O; the very definition of O precludes any such claim. If citizens are to all accept as free and equal persons a procedure to select from O, it must treat all their views of the common good in a fair way. But it is important that democracy also has a strong tendency to select from options within O rather than from outside of it. If we thought that democracy was fair but essentially random, selecting indiscriminately from O and $\neg O$ we may conclude that we are, after all, asses for being democrats. That is why no mere social equivalent of flipping the coin could serve as a political procedure; we need to do our collective best to identify O. This, no doubt, requires both public discussion and debate but also effective political institutions.

7.3 The First Lacuna in the Idea of a Fair and Efficient State

We want the state to choose from O and so efficiently promote the common good. Suppose the political procedure does not – it selects L_z, which is outside the optimal set. We know that we want to do the best we can to solve the complex optimization problem of the common good while recognizing that it is also inherently indeterminate. So we know that in this instance the state has failed to efficiently solve our problem. Has it also acted unjustly? After all getting stuck at a suboptimal arrangement is regrettable but need not be cause for deep political complaint. What would be such a cause for such complaint?

    We have thus far employed the idea of a perspective, which orders the options. Now it is a characteristic of an ordering that it is based on pairwise preference, and one can have a preference between really quite awful options.[35] If L_z outside of O is selected, this may be less efficient than options in O, but it nevertheless may reasonably promote the common good; on the other hand, it may be quite oppressive, and seriously set back the good of some citizens. Or, given her perspective, a citizen may simply think that some proposal is worse than no law at all on this matter (given some perspectives, one may think that *all* laws are worse than no law on this matter, since this may not be a proper matter for social regulation). But how are we to distinguish oppressive (or, we might say, strictly unacceptable) results from merely unhappy ones? Each citizen reflecting on various social worlds must distinguish (*i*) that which is optimal on her perspective (*ii*) those which are acceptable and (*iii*) those that that are strictly unacceptable. Certainly once we have abandoned the normalized version of the social contract one cannot reasonably equate (*i*) and (*ii*) – one cannot say that only

the optimal is acceptable – for such a stance precludes common acceptance of a body of norms or laws under conditions of diversity.[36]

Our question is what it means to say that a person vetos L_i. Recall Rawls's CI-procedure analysis of the universal law formulation of the categorical imperative (§2.1). Suppose one asks oneself whether one can will the altered social world that would result from L_i's introduction. It is important that one's concern is not simply that we all conform to L_i, but that each endorses the resulting social world – i.e., wills it. As the general will, one must will it – this means not only that one can will that others follow it, but that one accepts the law as one's own will. The point of thinking about the common good, after all, is for us to live a shared social life as free and equal persons. The aim of the social contract is for us to freely accept a common system of governance, in which, as Kant and Rousseau indicate, each is the legislator while each also obeys herself.[37] In more prosaic language, we can say that in this case the person internalizes the law's prescription: if confronted with the possibility of violating it without detection, and so with no fear of punishment or censure, she wills acting on it, and she will experience guilt should she fail to do so.[38] Our question is: given the way the common good maps on to some proposed law, does a person have sufficient reason to will it should it be selected?

This test is not arbitrary or contrived: only laws that can pass it are ones that each person can take as her will on this matter. If, when confronted with a possibility of undetected violation, she acts on her private will without guilt or remorse, a citizen does not will the law as having authority over her. Now a citizen may conclude that possible law L_i (which may or may not be a member of O, a set of optimal laws on this matter) does not adequately promote her view of the good. Given the costs of willing such a law, the good as defined by her perspective is set back by coming to will the law. For accepting a law as one's will has real costs: one obeys it even when one might better advance one's conception of the good by doing violating it, one feels guilt when one gives in to that temptation, and one must acknowledge that others may rightly take one to task, and even punish one, for such violations.[39] Once we allow a diversity of perspectives to come into play, we must accept different people will reach this decision at different points and for different reasons, but everyone will at some point come to it.

So we must divide each person's ordering into an eligible and ineligible set: those laws that are eligible are ones that, should the political procedure select them, she can will them; the ineligible ones, then, are those that she cannot. Let us call the socially eligible set E – the set of laws that are eligible in all citizens' rankings. We now see that a fair, efficient, *and non-oppressive* state will fairly select laws that, among those that are eligible, are not Pareto dominated by any other law. That is the goal of a political order aiming at the common good in a diverse society.

7.3 The Second Lacuna in the Idea of a Fair and Efficient State

Thus far we have been supposing that citizens agree on what constitutes a fair way of selecting from the *optimal eligible set*. But surely once we have introduced a diversity of perspectives this too is something about which we will disagree. As Nozick reminds us, "When sincere and good persons differ, we are prone to think they must accept some procedure to decide their differences, some procedure they both agree to be reliable and fair.... [But] this disagreement may extend all the way up the ladder of procedures."[40] No doubt once again we will be confronted with an optimal eligible set of political procedures or regimes: appealing to agreed-upon political procedures to extricate us from the indeterminacy of the diverse contract cannot solve *this* problem without begging the question. It would seem impossible for the political order to solve this problem, for this is the problem of what is the justified political order.

8. The General Will as a Social Phenomenon

The key to filling in this second lacuna is to understand – *pace* the social contract tradition – that we can share a general will apart from political determinations, and so we can have a social-but-not-political general will that identifies what political order, from the optimal eligible ones, we all will. To see how a general will can arise from a optimal eligible set without the appeal to political procedures, let us consider a case of selection of an informal social norm (or what we might call as a rule of social morality) from the optimal eligible set $\{N_1 \ldots N_n\}$.[41] Let us further conceive of each person having an overall "utility function" (a mathematical representation of the choice worthiness of social arrangements based on her perspective) as divided into two parts: (*i*) a part based solely on her ranking of $\{N_1 \ldots N_n\}$; she puts value on living according to a norm she ranks highly (she gets her preferred result) but (*ii*) she also values society converging on a common social norm (that there is a result; we all will a common norm). In the two-person case over $\{N_a, N_b\}$ case this yields an impure coordination game, as in Display 5.

		Betty	
		N_a	N_b
Alf	N_a	3 / 2	0 / 0
	N_b	0 / 0	2 / 3

DISLAY 5 COORDINATING ON A NORM

The numbers in the matrix refer to ordinal utility, with high numbers indicating highly ranked options. The uncoordinated outcomes indicate no shared norm. Looked at *ex ante*, Betty's perspective gives her reason to accept N_b over N_a; Alf's lead him to accept N_a over N_b. *Ex ante*, Betty does not have reason to accept N_a over N_b, nor does Alf have accept N_b rather than N_a. They do, however, have reason to coordinate on either of the norms rather than none at all (this is the importance of the second part of their utility function). Should Alf and Betty find themselves in a society characterized by N_a neither would have reason to change his or her action. Given each of their understanding of the good, they have the most reason to will N_a. Should they instead find themselves in a society characterized by N_b, each will then have most reason (given his or her conception of the good) to endorse it. Note that in neither case is any party induced by some external consideration to endorse a norm: consulting simply his or her own perspective, each has decisive reason to freely endorse whichever norm on which they have coordinated. In one society Betty can demand that Alf conform to N_b and, consulting only his own evaluative standards, he will have a reason to conform; in a society characterized by N_a Alf can demand that Betty conform to its demands and Betty will have sufficient reason to do so. And this even though, from the *ex ante* view, neither had reason to endorse the other's preferred norm. In this case, so long as N_a and N_b are both in the socially eligible set E, both are Nash equilibrium solutions.

Let us make the model more adequate. First, of course, we need to suppose that we have a large-numbers game. Second, we should allow that there will be great diversity in the weight different people put on the two parts of their utility function: some greatly care that the result is their preferred result, others have only a modest preference for one social world over another, but do care deeply that we all accept the normative arrangement. Now given these two steps towards realism, we are confronted with an increasing returns model.[42] In an impure coordination game with large numbers of people and multiple equilibria a bandwagon effect often takes over.[43] Suppose at time t_0 we have an absence of coordination on either N_a or N_b. Some play N_a and some play N_b; some greatly prefer N_a to N_b and vice versa, some have weaker preferences, and some are indifferent. Suppose that there is a chance event – perhaps simply the popular impression that N_a is more favored. In this case those with a weak preference for N_b over N_a, but a strong preference for coordination, will be apt to switch to N_a, as the best bet for coordination. Next those with a slightly stronger preference for N_b over N_a will observe that, despite their preference for N_b, an increasing number of the population are playing N_a; despite their preference for N_b, they too will see N_a as the better option for satisfying their coordination preference and their overall utility function. As more and more people switch to N_a, even those with a strong preference for

N_b will conclude that there is no chance of a general coordination on N_b; to insist on playing N_b would give them their 0 payoff from absence of coordination. The cascade towards N_a may continue until everyone, even those with a very weak preference to converge and a very strong preference for N_b over N_a switch to N_a as the only possible coordination point. Display 6 summarizes the dynamic.
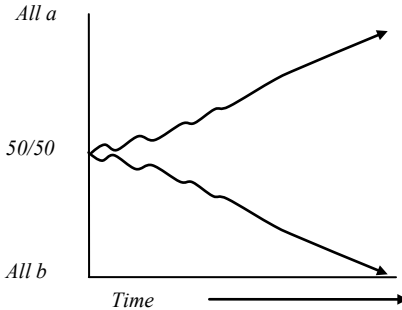
DISLAY 6 AN INCREASING RETURNS DYNAMIC[44]

As we can see, starting out with a population evenly split between advocates of norms *a* and world *b*, random events can lead the population to all N_a or all N_b equilibria. Which equilibrium emerges will be path-dependent: at time zero there is no reason why one or the other should emerge as the unanimously-selected choice. But once we have arrived at such a convergence, each citizen, consulting only her own evaluative standards, will freely accept the chosen social world. For our purposes what is crucial is that the contingent and accidental way in which large groups can come to coordinate on a common norm is no bar to there being a determinate general will that all can endorse given their different perspectives.

Note that this selection process is an actual collective social choice. In this sense, choice by a society takes up where discovery leaves off. To be sure, this is not an intentional "we-choice"; it is a social choice that arises out of a multiplicity of individual choices. Neither is it an abstract choice from some impartial Archimedean perspective outside our real social world. It is a collective choice that arises out of the social nature of individual choices: each person choosing to do what his perspective recommends given what others are doing.

On the account sketched here, the general will can be the outcome of a social process, partly a process of discovery and partly an outcome of social choice. Because this process is simply the outcome of each person acting on

her own perspective as she sees fit given what others are doing, the outcome genuinely expresses her will, though she has not bound herself to the outcome ahead of time. And that is why the process need not itself be justified, for we are not bound by the process, but by our own will in the light of the legitimate wills of others. It is this social process, which can lead to a uniquely justified outcome but which itself need not be justified, that allows society to form a general will on a variety of matters without resort to the political process – including the crucial question of what political process we all will. This primarily social rather than political interpretation of the general will was advanced by the British Hegelians. T.H. Green understood a society's morality as constituting a recognition of claims based on a common good.[45] Strikingly, Bernard Bosanquet understood the general will as the outcome of something very much like an invisible hand process, in which each individual, following her conception of the common good, helps to produce an overall social order and institutional structure willed by all.[46] What is important on this view is the way in which morality arises out of real social processes: it not a social contract in the form of an imaginary agreement, but a dynamic process of social discovery and choice that creates an actual social and moral fact – a social world that all will.

9. Conclusion

The idea of a social world, and moving from one to the other is, I admit, somewhat unusual in contemporary liberal political philosophy. But I believe it has a great benefit of helping us see that many of our disputes about social and political morality are not simply about the values relevant to social evaluation, but the social worlds that we evaluate. If we are to think dynamically about the general will and the common good, we must not simply consider what is best in our social world, but whether we should move to neighboring ones. An odd way to talk perhaps, but I think one that helps us see important problems more clearly.

I have tried to show that the normalization contract, in its search for consensus, prevents dynamic efficiency and discovery, being destined to endorsing only local optima. A general will that allows for discovery, looking beyond our current social world, must admit a diversity of perspectives. But to admit a diversity of perspectives implies, in turn, that there is no uniquely best general will. Taking seriously the general will as a process of discovery inevitably leads to also appreciating how it is a process of social choice. A true general will occupies the always-changing space between the discovery of the best way to arrange our common affairs and what, at any given time, all will.

ACKNOWLEDGEMENTS

NOTES AND REFERENCES

1. John Rawls, *Lectures on the History of Political Philosophy*, edited by Samuel Freeman (Cambridge, MA: Harvard University Press, 2007), p. 226.

2. Ibid., pp. 229ff.

3. Ibid., p. 224.

4. Ibid., p. 224.

5. Ryan Muldoon's *Diversity and the Social Contract* (Ph.D. dissertation, University of Pennsylvania) defends much the same conclusion, partly on similar grounds, though in crucial respects his fascinating analysis differs from that offered here.

6. On perspectives, see Scott Page, *The Difference* (Princeton: Princeton University Press, 2007), pp. 30ff.

7. As, for example, Amartya Sen endorses in his *Idea of Justice* (Cambridge, MA: Harvard University Press, 2009).

8. Thus, *pace* Sen, a contract theorist employing the idea of evaluating resulting social worlds will not be an "institutional fundamentalist" who is unconcerned with realizations, but rather a "comprehensive institutionalist" who is concerned with both procedures and their outcomes. I argue this in my "Social Contract and Social Choice," *Rutgers Law Journal*, forthcoming.

9. John Rawls, "Themes in Kant's Moral Philosophy," in *Collected Papers*, edited by Samuel Freeman (Cambridge, MA: Harvard University Press, 1999), p. 499. See also John Rawls, *Lectures on the History of Moral Philosophy*, edited by Barbara Herman (Cambridge: Harvard University Press, 2000), pp. 167–68.

10. Rawls, "Themes in Kant's Moral Philosophy," p. 500.

11. This is a matter on which different viewpoints disagree. See S.I. Benn and G.F. Gaus, eds. *Public and Private in Social Life* (New York: St. Martins, 1983).

12. See G. A. Cohen, *Rescuing Justice and Equality* (Cambridge: Harvard University Press, 2008), chap. 1.

13. Rawls, *Justice as Fairness: A Restatement*, edited by Erin Kelly (Cambridge, MA: Harvard University Press, 2001), p. 69.

14. Ibid., p. 70.

15. John Rawls, "The Idea Public Reason Revisited" in his *Collected Papers*, pp. 595ff.

16. Jürgen Habermas, "Reconciliation through Public Reason: Remarks on John Rawls's Liberalism," *Journal of Philosophy* 92 (March 1995): 109–131 at p. 129.

17. Rawls, *Justice as Fairness*, p. 166; Rawls, "The Idea of Public Reason Revisited," p. 599.

18. Perhaps we could only have such a condition under a sort of primitive anarchistic communism.

19. See Page, *The Difference*; Fred D'Agostino, "From the Organization to the Division of Cognitive Labor," *Politics, Philosophy & Economics* 8 (January 2009): 101–129 and his *Naturalizing Epistemology* (London: Palgrave, 2010). See also Michael Weisberg and Ryan Muldoon, "Epistemic Landscapes and the Division of Cognitive Labor," forthcoming *in Philosophy of Science*.

20. See D'Agostino, *Naturalizing Epistemology*, pp. 118–9.

21. We can depict the relation of these values as itself forming a rugged landscape; our aim is to search the terrain for the optimal combination. See D'Agostino, "From the Organization to the Division of Cognitive Labor;" Weisberg and Muldoon, "Epistemic Landscapes and the Division of Cognitive Labor."

22. Page, *The Difference*, p. 47.

23. See John Locke, "A Letter Concerning Toleration" *in Two Treatises of Government and A Letter Concerning Toleration*, edited by Ian Shapiro (New Haven: Yale University Press, 2003): 215–54 at pp. 245–6.

24. The metaphor of a rugged mountain range, with peaks and valleys leading up to the highest peak, has been employed in the analysis of ideal and nonideal theory in, for example, A. John Simmons's "Ideal and Nonideal Theory," *Philosophy and Public Affairs* 38 (Jan 2010). pp. 34–5. Compare Sen, *The Idea of Justice*, p. 102. It should be stressed, though, that the claim that pursuit of local optima may take one further away from the global optimum depends on the optimization problem having the characteristics of a rugged landscape. We should not get too carried away with metaphorical mountains and ignore the formal properties of the optimization problem. Perhaps more importantly, an uncritical use of the mountain metaphor implies that the rugged landscape is part of the world, rather than our perspective on the social world.

25. In the end, I believe that Rawls adopts this sort of partly normalized perspective in *Political Liberalism*, paperback edn. (New York: Columbia University Press, 1996), p. 225.

26. See David Schmidtz, "Non-Ideal Theory: What it is and What it Needs to Be." *Ethics*, vol. 121, (July 2011): 772-796.

27. A.K. Sen, "Maximization and the Act of Choice," in his *Rationality and Freedom* (Cambridge, MA: Harvard University Press, 2002), pp. 159–205.

28. See Robert Sugden, "Contractarianism and Norms," *Ethics* 100 (July 1990): 768–86.

29. Sen, "Maximization and the Act of Choice," p. 184. See also Kfir Lliaz and Efe A. Ok "Indifference or Indecisiveness? Choice-Theoretic Foundations of Incomplete Preferences." *Games and Economic Behavior* 56 (2006): 61–86.

30. For a classic paper on the problem, see Robert Aumann, "Utility Theory without the Completeness Axiom," *Econometrica* 30 (1962): 445–62.

31. I argue for this claim in "The Property Equilibrium in a Liberal Social Order (or How to Correct Our Moral Vision)," *Social Philosophy & Policy*, vol. 28 (Summer 2011): 74-101.

32. "A piece of legislation deserves respect because of the achievement it represents in the circumstances of politics: action-in-concert in the face of disagreement." Jeremy Waldron, *Law and Disagreement* (Oxford: Oxford University Press, 1999), p. 104. Waldron, however, fails to see that it must do so within an acceptable set. See my *Contemporary Theories of Liberalism: Public Reason as a Post-Enlightenment Project* (London: Sage, 2003), chap. 3.

33. I develop this conception of the state in more depth in *Justificatory Liberalism* (New York: Oxford University Press, 1996), Part III.

34. David Estlund, *Democratic Authority (*Princeton: Princeton University Press, 2008), esp. chap. 1.

35. This is one of the many reasons that it is mistaken to equate "preference" with "desire": one can abhor *x* less than *y*, but certainly desire neither, but in this case one still prefers *x* to *y*.

36. On the way in which insisting on one's optimal option precludes the possibility of solutions promoting a common good, see my "On Two Critics of Justificatory Liberalism," *Politics, Philosophy & Economics* 9 (May 2010): 177–212, pp. 200ff.

37. See Rousseau, *The Social Contract*, p. 176 (Book I, chap. vii); Immanuel Kant, *Foundations of the Metaphysics of Morals*, edited and translated by Lewis White Beck (Indianapolis: Bobbs-Merrill, 1959), p. 52 [Akademie, 433–434].

38. I provide a rather more subtle version of this test in *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World* (Cambridge: Cambridge University Press, 2011), chap. V.

39. As Kurt Baier would stress, compliance is everyone's business, see Kurt Baier, *The Moral Point of View*, abridged edn. (New York: Random House, 1965), pp. xviii–xix.

40. Robert Nozick, *Anarchy, State and Utopia* (New York: Basic Books, 1974), p. 98.

41. See P.F. Strawson, "Social Morality and Individual Ideal," *Philosophy* 36 (1961): 1–17.

42. The path-breaking work on increasing returns was done by W. Brian Arthur. See his *Increasing Returns and Path Dependency in the Economy* (Ann Arbor: University of Michigan Press, 1994). Those who are technically minded will see that this problem can be analysed in terms of positive externalities arising from network effects.

43. I am not claiming that convergence on a common norm always occurs; a polymorphic equilibrium may arise. See *The Order of Public Reason*, p. 398.

44. Adapted from Arthur, *Increasing Returns and Path Dependency in the Economy*, p. 3.

45. See T.H. Green, *Prolegomena to Ethics* (Oxford: Clarendon Press, 1890), Book III.

46. Bernard Bosanquet, *The Philosophical Theory of the State* in *The Philosophical Theory of the State and Related Essays*, edited by Gerald F. Gaus and William Sweet (Indianapolis: St. Augustine Press, 2001), and Gerald F. Gaus, "Bernard Bosanquet's Communitarian Defense of Economic Individualism" in *The New Liberalism: Reconciling Liberty and Community*, edited by David Weinstein and Avital Simhony (Cambridge: Cambridge University Press, 2001): 136–158.