Achieving a Realistic Utopia:

Rawls, Realization, and the Task of Political Philosophy

A thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF ARTS

in

THE DEPARTMENT OF PHILOSOPHY

of

RHODES UNIVERSITY

by

Rosa Elizabeth Terlazzo

November 2009

Abstract

In this thesis I argue that the tradition of political philosophy which follows in John Rawls's footsteps is obligated to concern itself not only with the realizability, but also with the realization, of justice. Although Rawls himself expresses a commitment only to the former of these, I argue that the roles which he assigns to political philosophy require him to take on the further commitment to realization. This is because these roles are meant to influence not only political philosophers, but the citizens of the wider community as well. The realistically utopian role, which I take to be the central one, requires political philosophy to inspire in that population a hope which I argue that realizability alone cannot provide. Given the deep revisions regarding the political nature of justice as fairness which Rawls made on the basis of realizability concerns, I argue that his theory must in this case be committed to a similar revision. The hope which political philosophy is meant to provide is simply not realizable until the discipline concerns itself centrally with the task of realization.

Acknowledgments

I would like to thank my supervisor, Tom Martin, for taking on my project, and for rolling so gracefully with the punches that came with its constant changes. Thanks to him for trusting my abilities when I didn't, and for reminding me of life beyond my thesis over all of those cups of coffee.

Thanks as well to the other members of the Rhodes Philosophy Department for being at once so demanding and so gentle with their postgraduate students.

The financial assistance from the Rhodes University Prestigious Scholarship toward this research is gratefully acknowledged. Opinions expressed and conclusions arrived at are my own and are not necessarily to be attributed to Rhodes University or the donor.

And thanks, finally, to Jonathan and the rest of my South African family, who gave this American a home that she never expected to find in Grahamstown.

Abbreviations

In-text references to John Rawls's works will use the following abbreviations:

JF	*Justice as Fairness*
KCMT	"Kantian Constructivism in Moral Theory"
LP	*The Law of Peoples*
PL	*Political Liberalism*
TJ	*A Theory of Justice*

Table of Contents

Introduction: Realizability and Realization in Political Philosophy....1
Chapter 1: Rawls's Work and its Commitment to Realizability....7
Ideal and nonideal theory....9
Ideal theory as thought experiment: the original position, the veil of ignorance....12
Ideal theory and its distance from our world: a well-ordered society....17
Nonideal theory: transition....19
Ideal theory and stability....20
Stability and realizability....24
Stability limited to a well-ordered society....26
Realizability and realization....27
Chapter 2: The Roles of Political Philosophy....29
The four central roles of political philosophy....29
The practical function....29
Orientation....31
Reconciliation....32
Political philosophy as realistically utopian....33
The import of these central roles....33
Realistic utopia....35
Realistic utopia is unlike other utopias....37
The ambiguity of realistic utopia....40
Realistic utopia is nonetheless problematic....44
Realistic utopia, realizability, and realization....46
Why Rawls is wrong to settle for realizability....47
Chapter 3: Hoping for a Realistic Utopia....51
Question 1: Can the deep tendencies and inclinations necessary for a well-ordered society obtain?....51
Question 2: Do we have reason to hope that the tendencies and inclinations necessary for a well-ordered society *will* ever obtain?....54
Our hopes have a causal relationship with the world....59
Hope aids us in more than fair gambles....60
Hope engenders new constitutive hopes....63
The pleasures of hoping....64
The shortcomings of Bovens's account of hope....67
Hope and epistemic irrationality....70
Some final words on hope and realistic utopia....71
Conclusion: What Comes Next?....74
Writing nonideal theory as transition....74
Implementing nonideal theory as transition....79
Bibliography....82

Introduction: Realizability and Realization in Political Philosophy

I have an intuition that I cannot shake: writing about justice means caring about injustice. When I read philosophical work on justice, part of me is pleased to have a better grasp of its constitutive parts, and of what an ideally just world might look like or entail. But a much, much larger part of me is dismayed that our own world doesn't look anything like that ideally just world that is being discussed. And the largest part of me by far is just *frustrated*. As philosophers who think about justice, we have a better sense than most of what our goal ought to look like. As even moderately aware and educated individuals, we can plainly see that what we have now is not what our philosophy is describing. And finally, simply as human beings, insofar as we wholeheartedly endorse a particular norm, we just *do* care about whether or not it obtains. When we as political philosophers recognize ourselves (as I think we cannot help but do) as those uniquely situated at the intersection of these three positions – philosophers who conceptualize ideals of justice, educated individuals who recognize our realities, and human beings who care in a deep way about the things which we take to have value – I can't help feeling that we really only have one option. We are in perhaps a better position than the members of any other single profession to conceptualize the world as it should be. If we also *care* about how the world should be, how can we not feel drawn to work in whatever immediate and practical ways that we can to bring about the world that we want? All of that thought, those hours, days, months, and even years thinking about what justice requires, and how the world ought to be – writing about justice just *must* mean caring deeply when justice is violated and abused.

Recently, a friend of a friend let me read his dissertation on the feasibility debate in philosophy (Ferriera 2009). In the simplest terms, the debate centers around the question of whether or not "ought" implies "can". Political philosophers tend to be of two minds when it comes to this question. One group thinks that the obligations placed upon us by principles of justice are in no way lessened if meeting them is beyond our power. In

their eyes, the demands of justice are unconditional, and "ought" therefore need not imply "can". The other group thinks that the only demands of justice which we can be obligated to abide by are those which we have the ability to meet. For them, demands of justice apply to humans, and it can only be true that we *ought* to do those things which are within our limited human power. In this case, feasibility is a strict limiting condition on the sorts of demands that justice can make.

My treatment of this debate has been, I recognize, remarkably brief. This is because it is simply not my concern here. At the back of my mind, as I read Ferriera's dissertation, there remained for me the persistent feeling that the most important question of all was not being asked as a part of the feasibility debate. The debate certainly addresses many of the questions that are raised by an intuition like mine: if writing about justice means caring about injustice, then it will also mean caring about what justice can and cannot require us to do. Questions about feasibility will indeed help us to see how justice and the world fit together. Ferriera's dissertation was my first substantive introduction to the feasibility debate, and it introduced me to many of the specific elements of Rawls's work which have proved to be so central to my own project. He alerted me to the importance of both realistic utopia and hope. Yet he considers them only in the context of feasibility. Since feasibility is his concern, this seems appropriate – yet I cannot shake the feeling that discussing these concepts in light of feasibility alone is not enough. There seems to me to be a further, and more important question, which the feasibility debate fails to address: how should the recommendations made in light of our conclusions about feasibility be played out in relationship to our world?

Indeed, I am not sure that the feasibility debate simply leaves this question unanswered – I think that its terminology actually *discourages* the asking. Discussing "feasibility" makes it seem as if the debate is (or at least could be) self-contained. Once we've answered the question of feasibility, there doesn't seem to be any obvious further question to be asked. For this reason, I prefer to talk about realizability. Realizability, like feasibility, can be debated as a possible limiting condition on demands of justice. It lets us ask all of the same questions about "ought" and "can", and it lets us fit justice and the world together in the same ways. Indeed, I take the content of "realizability" to be no

different from that of "feasibility". I choose the former term, however, because of the obvious bundle of terms that comes with it. A proposal may be "unrealizable", "realizable", or "realized". "Realized", however, has no counterpart when we are discussing feasibility. Here, a proposal may only be "infeasible" or "feasible". There is no third term which is necessarily attached to that proposal's implementation. I choose to discuss realizability rather than feasibility because the former term invites us to ask an obvious further question in a way that the latter does not.

It is this further question which is of such crucial and central importance to both my intuition and my project in this thesis: once we have answered the question of realizability, the obvious next question is that of realization. By realization, I mean the actual bringing about of the conditions of justice in the real world. Realizability, like feasibility, asks about what justice can hypothetically require of us as both individuals and societies. Realization, on the other hand, asks us to take on the further task of engaging with the world in order to make these requirements features of real human societies. The feasibility debate, then, is by no means unimportant to my project. Entertaining an intuition like mine, however, requires treating that debate as a starting point, rather than as a complete conversation in and of itself. In this thesis, I take for granted that realizable conceptions of justice are the only ones whose obligations we must take seriously: attempting the realization of any other sort of conception would be a fool's errand.

But poorly articulated intuitions and convictions, no matter how strongly held, hardly make for good philosophy. In this thesis I have tried instead to get at my intuition by addressing a collection of smaller questions which arise from it: What is political philosophy meant to do? Is it really just metaphysics of justice, or does it have a role beyond determining justice's nature? What should be the relationship between the limitations of the world and our ideals for that world? When our philosophy doesn't fit the world, how should we react? Should we revaluate and reformulate our ideals, or does such revaluation and reformulation of ideals compromise them too much? And in the end, again: *what is political philosophy meant to do?* Once these questions have been answered, I hope to show that it is the case that, on the basis of the commitments that

political philosophy already has, my intuition is an instructive one.

I've chosen to undertake this project in the context of the work of John Rawls for several reasons. First, and most obviously, it was Rawls, in *A Theory of Justice*, who refocused the attention of the philosophical community on questions of social justice. Without him, it is hard to say whether there would be space for a thesis of this sort at all. Second, much of the current literature on justice, even if no longer a direct reaction to his work, still owes Rawls a heavy debt. His theories and concepts laid the groundwork for much of what we do today. If I want to consider the purposes of political philosophy, then, it makes sense to return to the principles on which the modern discipline was originally based. Certainly, in order to do political philosophy, we need not accept all of Rawls's principles - but if we do not, we should at least have a reason why. Finally, I concentrate on Rawls because I think that he has chosen the right side in the feasibility debate: he is committed to articulating principles of justice which could effectively govern real human beings. It seems to me that only a theory of justice of this sort will be a valuable tool in addressing real questions of injustice.

Nevertheless, I think that there are significant problems with Rawls's theory of justice. Although it falls on what I take to be the right side of the feasibility debate, like that debate, it does not engage with the further question of realization. And on its own terms, I will argue that this is unacceptable. The central goal of this thesis will be to show that, if we commit ourselves to the purposes of political philosophy which Rawls proposes, then we as political philosophers must also be committed to doing the work of realization.

In Chapter One, I provide an introduction to parts of Rawls's work which will be required to understand the rest of the thesis. Here, I have several aims: first, I outline certain key concepts in Rawls's work, including ideal and nonideal theory and the various roles which each of these can have; political liberalism and its ties to stability and realizability; and the conditions of a well-ordered society. Second, I argue for my claim concerning Rawls's position in the feasibility debate. Third, I show that, although Rawls is committed enough to realizability to undertake major revisions to his theory, he does nothing either to engage with realization or to recommend that other political

philosophers should.

In Chapter Two, I turn explicitly to the roles that Rawls proposes for political philosophy, in order to show, contra Rawls, that they should commit us to realization rather than simple realizability. I argue that Rawls fails to take seriously the necessary connection which all of these roles have to the world outside of the philosophical community, and show how all of the previous roles should ultimately be taken to be in the service of the final one which Rawls proposes: that of realistically utopian philosophy. I end the chapter with the claim that these roles of political philosophy require us to ask two further questions of Rawls's theory. These are roughly, 1) Can the conditions for a well-ordered society obtain? and 2) Do we have reason to hope that these conditions *will* obtain? If these questions cannot be answered in the affirmative, then I think that Rawls is committed to rework his theory in the same way that other failures of realizability have required him to do.

In the final chapter, I attempt to provide answers to these two questions, and to show that Rawls's philosophy does not give us the tools which we require to meet the purposes that he himself sets out. The greatest of these is hope, and here I show that, on the basis of what hope is, it cannot be achieved in a practically useful, epistemically responsible way unless we require political philosophy to concern itself centrally with realization. This means that not only Rawls, but also we as political philosophers who follow in his footsteps, are obligated either to engage in the work of realization, or to reject and replace the purposes of political philosophy which he sets out. But I think that we cannot do the latter, for his purposes, and the hope which they inspire, are intimately related to the sorts of creatures that we human beings are.

This thesis, however, may seem somewhat hypocritical. Although I argue that political philosophy must engage itself in the concrete work of realization, I do not myself do this. I admit that I do not - but that sort of project is one for an entire discipline, not a single thesis. Nevertheless, in my conclusion, I gesture to some of the ways in which political philosophy can do that work. This work, I suggest, will have two parts. First, political philosophers will need to write a new sort of political philosophy: they will need to write what I call nonideal theory as transition. This sort of philosophy will engage

with questions of what we may acceptably do in our nonideal world that will bring us closer to the ideal world of a well-ordered society. Second, political philosophers must engage with members of other disciplines and industries to put these recommendations into practice. I will make particular suggestions for our educational and legal systems, but recognize that these are but two of the many ways in which political philosophy can join with other disciplines to do the work of realization.

This will mean maintaining a focus on the traditional subject matter of political philosophy: we will certainly need to continue to think about appropriate political ideals and goals. But it will also mean creating a new focus on a different sort of subject matter: the real world and the ways in which it can actually be changed will need to become our new focus. I think that it is appropriate to end this introduction with a quote from Barack Obama. He is not a philosopher, but as a politician he is the sort of individual with whom political philosophers will need to engage if they are to take realization seriously. He is speaking here of the specific problem of race relations in America, but his broader point, I think, should be extended to all of our work on justice:

> To think clearly about race, then, requires us to see the world on a split screen – to maintain in our sights the kind of America that we want while looking squarely at America as it is, to acknowledge the sins of our past and the challenges of the present without becoming trapped in cynicism or despair. This doesn't need to be political in the wrong way – not just compromise. It's looking at what we want *and* what we have, to figure out how to get from one to the other (Obama 2006: 233).

Chapter 1: Rawls's Work and its Commitment to Realizability

So much of philosophy follows in Plato's footsteps. Political philosophy, I think, is no exception. His *Republic* paints a picture in which the character of the ideally just city is a matter of abstract truth which stands in no meaningful relationship to the empirical facts of our world. The nature of that ideal city remains the same, whether or not real human beings in the real world could ever meet its demands and serve as its proper citizens, for the demands of that ideal are no less pressing if they are impossible to meet. As Plato says of that city, "its pattern is already there in heaven" and "the question of its present or future existence makes no difference" (Plato 1966: 592b). It seems to follow from this that the way that the world *ought to be* has little to do with the way that it *actually is*. The role that political philosophy should play for Plato, then, is that of discovering the truth of the ideal city's nature, rather than investigating the relationship between that nature and the state of our own world.

It is interesting to note that in the work of John Rawls, who has been quite possibly the most influential political philosopher of the last century, this concern with the ultimate true nature of justice is largely left behind. Instead, on the basis of his commitment to real justice in real human democracies, he turns his attention from the question of justice's truth to the question of its reasonableness.[1] This is belied by the way in which he phrases what he takes to be the fundamental question of political philosophy: "What is the most acceptable political conception of justice for specifying the fair terms of cooperation between citizens regarded as free and equal, and as both reasonable and rational, and as normal and fully cooperating members of a society over a complete life, from one generation to the next?" (JF 7-8). By asking about the "most acceptable" conception of justice, he makes it at once clear that he is not concerned with an

[1] This is because claims of moral truth can only be made from within particular comprehensive doctrines. For this reason, Rawls rejects moral principles which are *true* for more narrow moral political principles which can be endorsed by those who subscribe to all reasonable comprehensive doctrines. For more on comprehensive doctrines and truth, see *Political Liberalism* pp. xx and the latter sections of this chapter.

exposition of justice which is true at the metaphysical level, and that there are a number of conceptions which are at least prima facie plausible candidates for the solution to the question. It should be clear, simply from his phrasing of the question, that Rawls is not undertaking the Platonic project of describing the truth of justice as it is laid up in heaven.

By the end of his life, Rawls had become concerned with articulating principles of justice which could be the focus of an overlapping consensus in a politically liberal constitutional democracy. Since the individual members of this sort of plural society will always hold a variety of different beliefs about the good – a situation to which Rawls refers as "the fact of reasonable pluralism"[2] – any principle by which all can agree to be governed must necessarily abstain from making claims about matters of ultimate and metaphysical truth. A principle of justice cannot be discovered and then imposed upon such a society, for the content of any such principle would necessarily violate the reasonable commitments and beliefs of many of that society's members. Rather, to answer political philosophy's fundamental question in this sort of society, Rawls's concern must be with what works; he must worry about whether and how a conception of justice can govern real human beings who disagree about many aspects of their world.

Yet if Rawls's project is not an attempt to discover the abstract truth about the facts of justice, neither is it purely pragmatic in the Rortian sense – it is not an attempt to replace a concern with truth with a concern with the distinction between "descriptions of the world and of ourselves which are less useful and those which are more useful" (Rorty 1999: 27). Rorty's pragmatism takes as its starting point the world which we have, filled as it is with confusion, mistakes, and potential. His attempts to create a more just world begin with the problems with which we actually find ourselves confronted, and similarly try to form solutions from the building blocks of the world around us. While Rorty's project is concerned with understanding and working within the actual world in which we live, many of the aspects of our own actual world are irrelevant to Rawls. Although

[2] Of course, some conceptions of the good will also be *un*reasonable. These conceptions are ones which are willing to use political power to suppress other reasonable but conflicting conceptions of the good. Rawls's concern will be with containing rather than accommodating these unreasonable conceptions. For more on reasonable and unreasonable conceptions of the good, see *Political Liberalism*, Lecture II, Section 3.

Rawls is indeed concerned less with truth than with what can in fact govern human beings, his work nevertheless concentrates on governing human beings in a world that is quite unlike our own. Fully understanding this will require spelling out in some detail a number of Rawls's basic concepts, and the remainder of this chapter will be spent on this project. The broad point for now is that the relationship between the ideal and the actual in Rawls is a complex one. He is concerned neither with explicating a conception of justice which is ideal in the sense of Plato's ideal city, nor with confronting our actual world's problems of injustice as Rorty might be. Rawls's program is importantly different from both of these, and at this point it remains ambiguous to what extent the ideal and the actual can indeed be separated in his work.

Ideal and nonideal theory

When I said that Rawls concentrates on a world quite unlike our own, I meant that his principles of justice are conceived in light of and chosen to govern a society which lacks many of our own world's propensities to injustice. Rawls finds it useful to break his theory of justice into two parts: ideal and nonideal theory. In *A Theory of Justice*, he expresses the distinction in the following way:

> The intuitive idea is to split the theory of justice into two parts. The first or ideal part assumes strict compliance and works out the principles that characterize a well-ordered society under favorable circumstances. It develops the conception of a perfectly just basic structure and the corresponding duties and obligations of persons under the fixed constraints of human life. My main concern is with this part of the theory. Nonideal theory, the second part, is worked out after the ideal conception of justice has been chosen; only then do the parties ask which principles to adopt under less happy circumstances (TJ 216).

Very roughly, ideal theory is concerned with working out the principles of justice which would best govern a real society under the circumstances most hospitable to justice's correct functioning. It envisions a society in which citizens have both the will and the ability to support and abide by those principles. Nonideal theory fills the gaps left by ideal theory: it concerns itself with questions of partial compliance and of the transition from our own world to the world to which ideal theory would apply. As Rawls notes in

the quotation above, only ideal theory is central to his project of conceptualizing his theory of justice. Although he does offer very brief discussions of partial compliance theory in a number of his works, nonideal theory as transition receives hardly more than a mention of its existence. Rawls spends no time giving content to this concept.

As Rawls says, nonideal theory is written after and in light of ideal theory. For this reason, I will also begin my discussion of the two with ideal theory. However, we must qualify the rough conception given above of ideal theory's subject matter. Even if ideal theory is the appropriate place to begin, it will need to be tempered by nonideal theory in at least three ways. In the first case, while strict[3] compliance with the ideals of justice is perhaps a realistic goal, total compliance without exception is utopian in the pejorative sense. A society in which every citizen without fail both desires and is able to support and abide by those principles is entirely unrealistic. In any real world, there will always be isolated individuals who will be insufficiently motivated to act justly, and will attempt to take advantage of society in whatever way they can. Plato's ideally just city, whose existence in heaven does not depend on its existence in our world, could excusably demand not only strict, but total, compliance of this sort. Rawls's theory of justice, however, concerned as it is with governing real human beings, cannot ignore that there will be such breaches in compliance with justice. Nonideal theory, then, will be needed to address the questions which arise from these breaches, and must be prepared to provide insight in matters such as those of the penal system, compensatory justice, and civil disobedience (TJ 8).

Second, any society made up of real human beings must be a historical society, in which both unintentional historical contingencies and more intentional, systematic, historically unjust institutions may have left behind imperfect situations to which a well-ordered, ideally just society must answer (TJ 216). Although incidences of this first sort – unavoidable physical disability, for instance, or a past in which a series of not unjust interactions between individuals has led to a present in which the life chances of individual citizens are radically different – might not be matters of justice in themselves, responding to them certainly is. And, when we are presented with situations in which

[3] That is, if we take strict to mean broad and deep, rather than exceptionless.

disparate life chances *have* sprung from a history of unjust institutions (as, perhaps, in the case of the great differences in wealth and income between post-apartheid South Africa's black and white populations), attending to these inequalities is even more obviously a matter of justice. Once again, while Plato could ignore these sorts of considerations, overcoming them by fiat in his ideal city, Rawls cannot. Although they remain outside of the realm of ideal theory, Rawls must nevertheless recognize that they are issues with which any real society must grapple. However, since none of these questions, although realities in even the well-ordered society to which ideal theory applies, can reasonably be thought of as the sorts of questions of strict compliance which would be dealt with by ideal theory, they must also be answered by nonideal theory.

In the third case, it should be clear that our actual society is not one in which there is general consensus, let alone broad will on the part of many or most citizens to behave justly. In order for human beings to be effectively governed by the sort of justice that Rawls proposes, many things will need to change. Nonideal theory will also be necessary in our thoughts about a transition from our own society to the one to which ideal theory would apply. It is this second type of nonideal theory that receives the least attention in Rawls's work, and I will argue that it is in fact precisely this inattention that renders the Rawlsian relationship between the ideal and the actual not only complex but also problematic. This second type of nonideal theory will be discussed further below.

Before we explore what exactly it is that places the bulk of Rawls's work so squarely in the realm of ideal theory, however, let us note again what "ideal" will not connote when we use it to describe that theory. Although it follows on from the contrast with Plato's ideal city, it is worth noting explicitly that Rawls is not concerned with justice in a *perfect* society. Rawls is concerned with a conception of justice that can effectively govern real human beings. Human beings can certainly be far better in a plethora of ways than they are now, and Rawls's ideal theory is designed with many of these improvements in mind. Human beings shall never, however, *be perfect*. Although Rawls concentrates on ideal theory, he recognizes that it will always need to be complemented by nonideal theory if it is to lead to a conception of justice which can serve to govern real human beings even in their best conceivable state. Some of the

realities which separate even this idealized society from a state of Platonic perfection may be regrettable (like the necessity of a penal system in any real society), while others most certainly will not be (think here of the burdens of judgement,[4] and the free use of reason in a democratic society). Regrettable or not, however, realities they remain. And given all of these realities, Rawls's aim in ideal as well as nonideal theory is to formalize the principles of justice which would effectively govern not a *perfect* society, but the best one for which the realities of our world allow us to hope.

Ideal theory as thought experiment: the original position, the veil of ignorance

Having noted what it does not mean for Rawls's theory to be ideal, we must now explore further what it does mean. His theory of justice is ideal in at least two important senses. The first of these, discussed in this section, concerns his use of thought experiments and the implicit convictions of democracy which we must assume to obtain in a society in order for those thought experiments to yield the appropriate results. The second, addressed in the next section, concerns the nature of the society to which justice as fairness is meant to apply. I turn now to the first of these. Understanding this facet of ideality will require a broader understanding of some of the central features of Rawls's account. Let us begin with the concept of justice as fairness.

It is hardly controversial to say either that different principles of justice have held sway in different parts of the world and at different points in history, or that these conceptions have been heavily influenced by the predominant conceptions of the good and of human nature particular to those places and times. At various historical moments, accepted principles of justice have not been taken to apply, or been applied very differently, to (for example) slaves, women, children, and non-whites. Even today, conceptions of justice vary widely. The conception of justice found in a libertarian state will differ greatly from one found in a state arranged primarily around concerns of social welfare.

[4] The burdens of judgement are the numerous reasonable sources of disagreement to be found in a society in which the right of conscience is honored. For a full list of these burdens, see *Justice as Fairness* pp. 35-36.

Rawls himself is working from a liberal democratic position which sees all human beings as moral persons, regardless of the morally irrelevant characteristics like race or gender which might be attached to them. The moral person is defined by two powers: first, she possesses the capacity for a conception of the good, which means that she can choose for herself a rational plan of life. Second, she has the capacity for a sense of justice, which expresses itself as a "regulative desire to act upon certain principles of right" (TJ 491). Since these moral powers are possessed by all members of a society,[5] it is equally important that each of these legitimate plans of the good be accommodated – albeit without any one of them inhibiting the plan of another. Given this, Rawls believes that the most appropriate conception of justice for governing such a society is a social contract conception to which all of its members would willingly agree. In his words, the appropriate principles of justice are "the principles that free and rational persons concerned to further their own interests would accept in an initial position of equality as defining the fundamental terms of their association" (TJ 10).

However, determining such principles will not be simply a matter of aggregating the preferences of the parties concerned. While Rawls's broad concern is to provide an answer to the fundamental question of political philosophy, his method is to explicitly contrast his justice as fairness with utilitarianism, which he takes to have been the day's prevailing conception of justice (TJ 3). He rejects utilitarianism for two reasons. First, it gives insufficient attention to the importance of the individual. Although utilitarianism is egalitarian in that it places no more value on the preferences of any one person than those of any other, it is insufficiently attentive to the individual in that it will willingly subsume the good of some smaller group of individuals to the good of a larger group if the resulting balance of preferences is sufficiently high.

A second ground for Rawls's rejection of utilitarianism applies to, but is hardly

[5] With a notable exception: adults who have (for example) suffered brain damage or experience severe forms of mental retardation may never be able to develop these capacities. Here, however, is yet another – quite interesting, although for my purposes far less central – example of the ideal character of Rawls's theory. In his work, he brackets the question of mental and physical handicap, proceeding instead as if all citizens were within a certain normal range of physical and mental ability (See, for instance, TJ 83-84). Clearly Rawls appreciates that this will not actually be the case in a real society, but he leaves the question of these handicaps for nonideal theory. For some remarks on why this bracketing might nevertheless be problematic, see Farrelly 2007.

exclusive to, that doctrine. All other things being equal, when choosing the principles that govern our lives, each of us would usually prefer a principle that benefited us more to one that benefited us less. In a utilitarian system, the right or the justice of a course of action will depend upon the likelihood that it will bring about the greatest aggregate happiness of those involved. The happiness of each individual will be based upon the preferences of that individual, which will themselves be based upon the position which that individual occupies in society. A principle of justice chosen in this situation, then, will be quite sensitive to the self-interested motives of those who would prefer to see their own position in society benefited.

In justice as fairness, Rawls seeks to correct both of these utilitarian shortcomings: its insufficient attention to the individual, and its willingness to make principles of justice contingent upon the societal positions of those choosing them. To counteract the first of these problems, he stipulates that the principles of justice must be ones to which not simply the majority, but *all* of the members of the contract party can reasonably agree. Given Rawls's conception of the person, he has already accepted a democratic regime as his basic starting point. However, his conception of the person entails further that an acceptable democracy cannot be a simple case of majority rule. Rather, it must balance equality and liberty so that each individual in that society retains the proper space to exercise her two moral powers (her respective capacities for a conception of the good and a sense of justice). Rawls holds that these values are implicit in the public culture of any functioning democracy, and that the best way of explicitly working them out is to see society as a fair system of social cooperation over time (JF 5). In such a society, citizens do not merely allow others to pursue their own plans of the good – by engaging in social cooperation, they actually *enhance* the ability of others to pursue these plans, as they enhance their own at the same time.

To address the second short-coming of utilitarianism, Rawls conceives of a thought experiment to help the members of the contract party to choose principles of justice which are not colored by their societal positioning and so are appropriate to society as a fair system of cooperation. This thought experiment is comprised of two major parts: the original position and the veil of ignorance.

Rawls adopts the convention of the original position because he is concerned that the principles of justice should be chosen by individuals who are equal parties to the contract.[6] If some members were to hold a position of high prestige, or to wield great economic power, then those members' contributions to the discussion could well hold disproportionate weight. We should note that this disproportionate weight need not be coercive for it to distort the agreement reached in the original position: since all of the parties involved in the contract discussion are thought to possess the two moral powers, it will not do for some members to be silenced out of awe or respect any more than out of fear. Principles meant to regulate society from now into the future are too important to be affected by "contingent historical advantages and accidental influences" (JF 16). To avoid these contingencies, Rawls imagines the principles of justice to be decided upon in an ahistorical "original position" in which the parties to the contract are situated outside of society. The original position, then, is meant to ensure the equality of the parties at the time of their decision upon the principles of justice.

However, there is also a more forward-looking reason to remove the question of societal position from the forming of the contract. Although parties in the original position are situated outside of society in order to choose the principles of justice, they are nevertheless choosing for the society in which they themselves would live, and they recognize that in that society they would hold a determinate social position. However, just as we cannot have certain parties wielding excessive influence in the original position, we can also not have parties motivating for principles of justice which would disproportionately benefit them in the determinate positions which they will eventually fill. For this reason, Rawls introduces a "veil of ignorance" which prevents the representatives in the original position from knowing the specifics of the society that they will enter or their position in it (TJ 118). Fairness is, for Rawls, one of the convictions implicit in the public culture of a democracy: we know that "the fact that we occupy a

[6] Recall that Rawls emphasizes the free and rational nature of the parties in a way that other contract positions do not. Consider the contrast, for example, between the social contract of Rawls and that of Hobbes, in which the equality of the parties is immaterial, and the justice of the contract is based upon its continuing ability to effectively shield its member parties from the state of nature. For more on Hobbes's social contract, see *Leviathan*. Rawls contract is novel, then, in that it is only just if it results from the uncoerced assent of those who stand on an equal footing with one another.

particular social position, say, is not a good reason for us to accept, or to expect others to accept, a conception of justice that favors those in that position" (JF 18). Nevertheless, he also recognizes that, while we may affirm this conviction abstractly, it is rather harder to live our lives in this way. In strict matter of fact, we must "nullify the effects of specific contingencies which put men at odds and tempt them to exploit social and natural circumstances to their own advantage" (TJ 118). The veil of ignorance, then, is implemented in order to remove all temptation to stack the principles of justice in favor of some members of society. Since each party behind the veil of ignorance realizes that she could enter into society in either the least or most advantaged position, she will want the principles of justice to promote fairness and the ability to pursue a rational plan of life in *whichever* position she might take up. In contrast to the original position, which ensures that past inequalities will not affect the principles of justice chosen, the veil of ignorance ensures that considerations of future disparities in resources and social position will not either cause the chosen principles of justice to disproportionately benefit one segment of society.

In the end, Rawls offers the following principles as those which would be reached by parties behind a veil of ignorance in the original position who are free and rational agents and are searching for the best principles to govern society as a fair system of cooperation over time:

> (1) Each person has the same indefeasible claim to a fully adequate scheme of equal basic liberties, which scheme is compatible with the same scheme of liberties for all; and
> (2) Social and economic inequalities are to satisfy two conditions: first, they are to be attached to offices and positions open to all under conditions of fair equality of opportunity; and second, they are to be to the greatest benefit of the least-advantaged members of society (JF 42-43).

So we should recognize that the principles of justice and the way in which they are arrived at are ideal in at least this way. The original position and the veil of ignorance, which yield the principles that Rawls endorses, are both thought experiments, which, while they are useful ways of thinking about the generation of principles of justice, are not positions into which we could actually enter. While they may be appropriate given

the normative picture painted of the implicit convictions of a democracy's public culture**,** there is some question as to how deeply and widely these convictions hold in the democracies of our actual world. Rawls's theory is ideal in this sense in that choosing its principles requires that we accept convictions which may not effectively motivate us in the actual world in which we live. Choosing as we are now, these convictions may well be absent, and our resulting principles are likely to look very different – perhaps more like the utilitarian principles which Rawls hopes to replace. Although this sort of ideality of theory is not the one that primarily concerns Rawls (to this second sort we will turn in the next section), it is one that we must bear centrally in mind when, in Chapter 3, we consider the tension between Rawls's emphasis on ideal theory and his emphasis on hope.

Ideal theory and its distance from our world: a well-ordered society

The sort of ideal theory on which Rawls concentrates is the one which applies to the persons and society which are to be governed by justice as fairness. The principles of justice are ideal in that they are meant to apply only to a well-ordered society. When Rawls first presents the idea of a well-ordered society, he stipulates that it must meet the following conditions: a well-ordered society must be governed by a set of rules designed to advance the good of its members, and it must be effectively regulated by a public sense of justice (TJ 4). This regulation has two sub-conditions: first, everyone in the society must accept and know that others accept the same conception of justice. Second, the basic social institutions of that society must generally satisfy – and generally be known to satisfy – the conception of justice publicly accepted by its people.

We should note, however, that these conditions for a well-ordered society do not themselves entail the sorts of principles of justice which Rawls endorses. Consider first that the rules which govern a society must be designed to advance the good of all of its members. Although the good concerned must be the good of all of those who take part in the system, this condition alone does not preclude a society governed by a heavy paternalism. It seems that a benevolent dictatorship could quite conceivably advance the

good of the people, if good were understood without reference to autonomy or its related concepts. But clearly this would not be acceptable in a constitutional democratic society, which is the kind of society which Rawls's justice as fairness is aimed to govern. The same complaint can be made about effective regulation by a public sense of justice. Here as well, a paternalistic but benevolent authority could dictate the terms of justice which governed a society without its "well-ordered" status being violated in any way. The point is that, although Rawls's theory of justice only applies to a well-ordered society, neither of the conditions for a well-ordered society nor the two conditions taken together presuppose or necessitate that a well-ordered society be the sort of society which *would* be governed by Rawls's theory of justice – in other words, a well-ordered society need not be a democracy. This need not be problematic, however. It is the content of the conception of justice in question which will limit the sort of governmental structure acceptable in a well-ordered society. In this case it is only necessary for ideal theory that the principles of justice as fairness apply to, not that they spring from, a well-ordered society.

For a well-ordered society to be the sort of democracy which *could* be governed by Rawls's theory of justice, at least two stipulations must be made. First, it must *be* a democracy, and not some sort of authoritarian state. This is because personhood for Rawls is defined by possession of the two moral powers.[7] Any system which refused citizens a role in governmental decision making would deny the importance of allowing persons the chance to exercise the first of these – their capacity to form a plan of the good. Second, the condition of effective regulation would have to be somewhat thicker. Here, it would require something more than embodiment by institutions and acceptance by citizens of the conception of justice. "Acceptance" must be made into a thicker concept which includes the *right sort* of motivation to be governed by that conception.[8] And indeed, later in *A Theory of Justice,* Rawls clarifies that societal regulation by a

[7] Or at the very least some relationship to the *capacity for* these powers. Requiring possession of the powers in order to attribute personhood raises all sorts of complex questions about the personhood of young children, individuals with brain damage, etc.

[8] Motivation by force or brainwashing, or motivation only in situations where complying with the conception furthers one's own self interests, for example, are examples of motivation for the wrong reasons.

public conception "implies that its members have a strong and normally effective desire to act as the principles of justice require" (TJ 398). A thinner version of acceptance could allow for a significant number of citizens to publicly accept the principles of justice as both reasonable and rational, but still recognize that violating those principles would prove personally beneficial in many instances. In this case, the society in question would be left with a significant free-rider problem, and any attribution of "well-ordered" status would seem ridiculous.

Clearly, the conditions of the well-ordered society to which Rawls's principles of justice are meant to apply are highly idealized. Our concern in this case is no longer that the principles of justice are reached only via a thought experiment into which we could not in practice enter – our concern is now that our world is lacking many of the characteristics that would make these principles effectively regulative. First of all, it is not the case that all members of our society accept the same conception of justice. As we noted before, a libertarian and a social democrat will have very different ideas of what appropriate justice ought to be. Second, even if everyone were to accept the same conception, people – a great number of them, if not perhaps the majority – may well not be sufficiently motivated to act on the corresponding principles. And third, since we do not have a single publicly shared conception of justice, it should go without saying that our basic social institutions do not satisfy the requirements that any such conception might have. Quite simply, our actual world is not one in which well-ordered societies exist. This sense of ideality, and the distance between a society to which justice as fairness could apply and our own, must also be central to our thoughts about the tension between ideal theory and hope in Rawls's work.

Nonideal theory: transition

Since well-ordered societies – the primary subjects of Rawls's ideal theory – do not exist in our actual world, let us turn to the sort of theory that concerns itself with the relationship between that sort of ideal world and our own. This is the second form of nonideal theory: that of transition. For Rawls, although ideal theory does not apply to the

world in which we live, ideal theory is nevertheless a prerequisite for the nonideal theory which might so apply:

> A conception of justice must specify the requisite structural principles and point to the overall direction of political action. In the absence of such an ideal form for background institutions, there is no rational basis for continually adjusting the social process so as to preserve background justice, nor for eliminating existing injustice. This ideal theory, which defines a perfectly just basic structure, is a necessary compliment to nonideal theory without which the desire for change lacks an aim (PL 285).

Nonideal theory as transition, then, is the working out of how we might transform our own unjust world by finding "policies and courses of action that are morally permissible and politically possible as well as likely to be effective" (LP 89). Unlike nonideal theory as partial compliance theory, this sort of nonideal theory as transition ought ultimately to make itself redundant; when we have attained (as closely as possible) the sort of just basic structure that ideal theory defines, the questions of nonideal theory as transition will be much less central, if not altogether irrelevant. However, for Rawls, we cannot coherently undertake the sort of project that nonideal theory as transition entails without first setting for ourselves a goal worth reaching.[9] Setting this goal is the work of ideal theory, and the necessity of the goal serves as Rawls's justification for concentrating his work in the ideal realm to the detriment of the nonideal.

Ideal theory and stability

At this point, we should have seen enough to know that Rawls's ideal theory at its most basic level takes very seriously the facts of at least some real world. From the very beginning, his ideal theory took into account the facts of human nature, the need for corresponding partial compliance theory, and the other general limiting facts of human societies. However, as his work progresses onwards from *A Theory of Justice*, he begins

[9] Of course, this will also be the case for nonideal theory as partial compliance theory. However, this modeling of ideal theory will effect the two kinds of nonideal theory in different ways. Nonideal theory as transition will take its goal from ideal theory in order to make itself redundant. Nonideal theory as partial compliance theory will take its goal from ideal theory in order to serve as a damage control mechanism which will never stop being necessary.

to take these constraints even more seriously. By 1995, when he wrote *Political Liberalism*, he had significantly reworked justice as fairness to account for the importance of the real-world condition of stability.

It is in *Political Liberalism* that Rawls first explicitly recognizes that democratic societies are centrally characterized by what he calls "the fact of reasonable pluralism" (PL 36). This is the fact that within any society in which equal liberty of conscience is recognized and affirmed, individual citizens are bound to call upon different religious, philosophical, and moral doctrines when they attempt to formulate their separate rational plans of the good. When these doctrines are comprehensive – that is, when they go beyond considering questions of the good narrowly applicable to the shared political realm, and instead include "conceptions of what is of value in human life, and ideals of personal character, as well as ideals of familial and associational relationships, and much else of what is to inform our conduct" (PL 13) – they are bound to conflict with one another. There are simply too many legitimate ways of answering these questions for all of us to settle finally upon the same one. So long as I affirm one particular comprehensive doctrine as true, I will be unable to affirm at least several of the other conflicting comprehensive doctrines that I find around me as similarly true.

The fact of reasonable pluralism becomes problematic when we recognize the importance of stability to any political conception. A just regime that could not sustain itself and would pass immediately away would provide a poor answer for specifying, as the fundamental question of political philosophy asks us to do, the terms of fair cooperation between free and equal citizens. If a conception is really to be the most appropriate for this task, it must also answer to the further condition: that such a scheme be the most appropriate *over time*. In other words, the most appropriate conception must be one which is *stable* over time. In Rawls's words, this means that "in order to be stable, a political conception of justice must generate its own support and the institutions to which it leads must be self-supporting" (JF 125).

With this new consideration in mind, Rawls changes his focus from the fundamental question of political philosophy to the fundamental question of political liberalism: "How is it possible for there to exist over time a just and stable society of free

and equal citizens, who remain profoundly divided by reasonable religious, philosophical, and moral doctrines?" (PL 4). In answering this question, we must take into account both the fact of reasonable pluralism and the importance that Rawls places on stability, and we should see quite quickly that a democratic society could not be stably governed by any one comprehensive doctrine. Since each possible comprehensive doctrine conflicts with several others, no single one of them could satisfactorily govern all citizens. And since there would be so much dissent among the population, no single comprehensive doctrine could generate its own support broadly enough to sustain itself and remain stable.

For this reason, having recognized that his own justice as fairness as presented in *A Theory of Justice* was comprehensive in this way, in *Political Liberalism* Rawls rejects that comprehensive doctrine for a narrower new *political* doctrine of justice as fairness. This political conception has three elements: first, a political conception applies only to the "basic structure" of society – to its "main political, social, and economic institutions, and how they fit together into one unified system of social cooperation from one generation to the next" (PL 11). In other words, the political conception only regulates the basic structure which assures citizens the rights, liberties, and basic goods that they need to pursue their own private projects and conceptions of the good. The political conception of justice is not instructive when it comes to the directions that these private projects themselves ought to take.

The second and third features of a political conception of justice go together. In the second case, the political conception is not derived from any one comprehensive view, or from a mixture of several. Instead, it is a freestanding view which presents the principles of justice by which any rational and reasonable[10] person should choose in the original position to be governed. Any reasonable comprehensive doctrine ought to be compatible with such a political conception. This is because a comprehensive doctrine which is reasonable must by definition recognize equal liberty of conscience and the

[10] These terms are importantly different, though related. A rational actor is one who acts to advance her own good. A reasonable person, on the other hand, recognizes the similar claim of others to pursue their *own* good, and so abides by the rules of justice and the fair terms of cooperation to which all would agree, even when doing so is to the detriment of her own interests (JF 6-7).

prerogative of other citizens to pursue their own rational plans of the good.[11] The political conception, however, is able to accommodate all of these reasonable comprehensive doctrines because it does not try to answer, as those various doctrines might, all questions of value and of human good in society. The third, and complimentary characteristic of a political conception of justice is that its content is derived from, rather than foreign to, the public political culture of a democracy: "In a democratic society, there is a tradition of democratic thought, the content of which is at least familiar and intelligible to the educated common sense of citizens generally" (PL 14). These three features, taken together, make it a conception that can gain the support of an overlapping consensus between the reasonable comprehensive doctrines in question. Since it comes from none of these comprehensive doctrines but is drawn from ideas familiar to all, and since it applies only to the basic structure which makes the expression of all of these doctrines possible, all comprehensive doctrines which are reasonable should be happy to affirm the political doctrine.

It is perhaps conceivable that we could create some sort of stability in a society without relying upon a political conception of justice – but we could not do so in a *democratic* society. If stability under one comprehensive doctrine could be possible at all, it would be so only through the violent and forceful imposition of that doctrine upon the people. Rawls calls this "the fact of oppression" (PL 37). (Imagine, for example, fascist Germany or the Spain of the Inquisition.) Although such "stable governing" might be possible, the factors for its possibility clearly preclude democratic government and equal liberty of conscience.

Even if, by some fluke or accident, a society could exist which could be governed by a single comprehensive doctrine (such as justice as fairness in its original form as a comprehensive view) which all shared, this should not be our ultimate hope. There are numerous sources of reasonable disagreement (which Rawls calls "the burdens of judgment"[12]) about the facts of even the deepest and most important matters that we must consider in our lives. If citizens are to consider for themselves the issues that they

[11] Recall that some comprehensive doctrines will not be reasonable ones. See footnote 2.
[12] For a full list of the burdens of judgment, see pages 35-36 in *Justice as Fairness*.

encounter, in the service of their own pursuits and plans of the good, rather than rely upon some authority, then these burdens of judgment will invariably lead to reasonable pluralism. As Rawls says, reasonable pluralism is "the natural outcome of the activities of human reason under enduring free institutions. To see reasonable pluralism as a disaster is to see the exercise of reason under the conditions of freedom itself as a disaster" (PL xxiv). And in a society in which the two moral powers[13] are the constituent parts of the normative conception of the person, such an outcome can hardly be seen as a disaster.

If stability is so important in the selection of a conception of justice, then the fact of reasonable pluralism demands that we choose a conception which people can in fact support. In the end, stability is a make-it or break-it condition for a Rawlsian conception of justice; if the principles chosen in the original position do not in fact prove to lead to stability when they are applied to a real society, then the project of choosing in the original position must begin anew. Since the facts of the real world would not allow justice as fairness as a comprehensive doctrine on these grounds, the facts of the real world and the importance of stability there necessitate its transformation into a political conception.

Stability and realizability

We have seen that stability is important enough to Rawls to cause him to rework one of the most central components of his theory of justice. But what exactly is the place of stability in Rawls's work? Although he does not explicitly say so, I take it that Rawls's emphasis on stability is part of his answer to the feasibility debate. This is because stability, for Rawls, works in the service of realizability.[14] Rawls's reason for rejecting his comprehensive doctrine of justice as fairness in favor of a political doctrine is that "as used in *Theory*, the idea of a well-ordered society of justice as fairness is unrealistic. This is because it is inconsistent with realizing its own principles under the best of foreseeable conditions" (PL xvii). In other words, the original version of a well-ordered society,

[13] Again, the capacity for a sense of the good and the capacity for a sense of justice.
[14] Recall from the introduction that I use feasibility and realizability interchangeably.

based as it was in a comprehensive doctrine of justice as fairness, is not the sort of thing that we *can* have in any sort of real world – because it cannot be stable, it cannot be realizable.[15]

Of course, it is not strictly true that in order to be realized a concept must also be stable. An unstable concept could perhaps be realized for a short time and then pass swiftly from existence. But this sort of concept will not be realizable in the way in which it must be to answer the fundamental question of political liberalism. Recall that this question requires us to explain "how it is possible for there to exist *over time* a just and stable society of free and equal citizens, who remain profoundly divided by reasonable religious, philosophical, and moral doctrines" (PL 4, my italics). So, while realizability may be a feature of many different sorts of conceptions, political and otherwise, the term "realizability" as used in this thesis will only refer to a feature of political conceptions which allows them to effectively answer this question.

It is clear, then, that realizability is a central and nonnegotiable condition for an acceptable Rawlsian theory of justice. If Rawls rejected his own theory because of questions of stability which were themselves fueled by the concern of realizability, then we must be open to the idea that other failures of realizability in his final conception would similarly call for major reworking of the theory.

Presumably, there are numerous ways in which a proposed society could fail to be realizable. However, I will be concerned with only one of them here. Recall that for the conditions of a well-ordered society to be met, it must be the case that its "members have a strong and normally effective desire to act as the principles of justice require" (TJ 398). Recall further that to meet the stability requirement, that effective desire to act must reinforce itself over time: it must "generate its own support and the institutions to which it leads must be self-supporting" (JF 125). Since stability is a requirement of realizability, a conception of justice which did not meet stability's own requirements would, on Rawlsian grounds, be unacceptable. If justice as fairness, then, were not able to meet the stability condition, then it would be unrealizable as well, and would need to be reworked accordingly.

[15] Stability, then, will be a necessary, but not a sufficient, condition for realizability.

I think that, if we were to apply these two requirements to our own actual world, they would not be met. It seems uncontroversial to claim that at least the majority of our worlds' citizens do not have either a strong or normally effective desire to act on the principles of justice, and if this sort of inclination on the part of the population does exist, then it is clear that the population will not create self-supporting just institutions. Applied to a current real-world society, then, justice as fairness would fail to meet the stability and realizability conditions, and so would need to be reworked.

Stability limited to a well-ordered society

But Rawls is not concerned with our real-world society. He has considered the possibility of a population in which the appropriate sense of justice obtains, and he feels that it is not unrealistic to imagine that this sort of society might obtain. After all, he is working with an idealized conception of human beings and of the basic structure which would influence their formative years. According to him, if citizens grow up in a basically just society, then they ought to automatically acquire the desire to be governed by the appropriate conception of justice:

> The idea is that, given certain assumptions specifying a reasonable human psychology and the normal conditions of human life, those who grow up under just basic institutions – institutions that justice as fairness itself enjoins – acquire a reasoned and informed allegiance to those institutions sufficient to render them stable. Put another way: citizens' sense of justice, given their character and interests as formed by living under a just basic structure, is strong enough to resist the normal tendencies to injustice. Citizens act willingly to give one another justice over time. Stability is secured by sufficient motivation of the appropriate kind acquired under just institutions. (JF 185).

Since Rawls's philosophical project is only concerned with ideal theory, it seems perhaps acceptable for him to make that leap. If (as would be the case in the sort of society described by his ideal theory) individuals were raised by parents who wholeheartedly endorsed and were sincerely motivated by the common sense of justice which was shared in their community and embodied in their institutions, then those individuals may well be similarly motivated to give each other justice and avoid injustice. Since the emphasis in

Rawls's work is on whether justice as fairness could regulate a well-ordered society, and not on whether it could regulate our own, the question of stability in our current circumstances need not be an issue for him. It follows from this that, if his assumptions about the possibility of this world of ideal theory are correct, he has met the realizability condition as far as he needs to.

Realizability and realization

I am skeptical, however, that realizability ought to be where Rawls's concern with the real world ends. There is, after all, the further issue of realization. Thomas Pogge, in a discussion of Rawls, suggests that "the point of political philosophy is not merely to show that certain principles are true, but to make them true by motivating the struggle for their gradual implementation" (Pogge 1989: 6). This claim seems to me to be correct, at least in the context of a certain kind of political philosophy. Returning to Plato's ideally just city, it is clear that this kind of political philosophy is unconcerned with either the possibility or the project of motivating much change in the real world. In other words, it is unconcerned with both realizability *and* realization. Rawls's own sort of political philosophy, however, is concerned at least with realizability – and in order to ask whether it ought also to be concerned with realization, I think that we must ask ourselves just how this concern with realizability is motivated.

Although Rawls's work is primarily in ideal theory, as we have seen, he nevertheless recognizes in at least some sense that there is another vital part to our political philosophy. He says that "the task is to articulate a public conception of justice that all can live with who regard their person and their relation to society in a certain way. And though doing this may involve settling theoretical difficulties, the practical task is primary" (KCMT 519). This is not a question of discovering truth – it is a very practical and concrete question about the best ways in which real people really *can* live their lives.

It is well and good, then, to say that ideal theory provides the appropriate aim for nonideal theory, and that the former must be worked out before we can consider the latter. But this does not seem to me to be enough. Although Nicholas Ferriera's primary

concern is determining feasibility (rather than bringing about the realization of principles of justice, as mine is) I think that he is right to question the sort of stability that Rawls makes his theory's acceptability conditional upon. Ferriera thinks that Rawls pays too high a price when he makes such serious modifications to justice as fairness on the basis of the sorts of assurances about stability that ideal theory can give us. Says Ferriera, "If stability is desirable in politics, it seems at least plausible that only actual stability is desirable, especially if substantial costs must be incurred to obtain it" (Ferriera 86). It seems to me right to say that, if we turn our thoughts on justice to the sort of justice that people *could* agree to be governed by, we should also turn them to the sort of justice that people *would* agree to be governed by, and that we should concern ourselves as well with seeing that we do what we can to bring such a conception of justice about.

Indeed, although Rawls's work remains primarily in ideal theory, I do not think that he would do well to content himself with realizability in an ideal society, and to leave aside questions both of realizability and realization in our own world. In fact, on Rawls's own terms, I think that there is great reason for him (and generally for the strand of political philosophy which carries on his legacy) to attend to these problems. To defend this claim, however, I will need to examine the central roles that Rawls proposes for political philosophy. Understanding why these roles entail that concerns of realizability must lead us to further concerns of realization will be the task of the next two chapters.

Chapter 2: The Roles of Political Philosophy

At the end of the last chapter I suggested that if we are to take Rawls's theory seriously, then we as political philosophers should be as concerned with realization as we are with realizability. Although, as we shall see, Rawls maintains until the end that ideal theory, and its corresponding concern with realizability, should appropriately remain the subject of his political philosophy, I think that his later work only gives us increasing reason to believe precisely the opposite. There more than ever it should become clear that meeting the conditions of realizability cannot be the final aim of political philosophy.

The four central roles of political philosophy

Rawls proposed his theory of justice as an attempt to remedy the shortcomings of the political philosophy that had come before it. He recognized the great need for revision in the discipline, but recognized as well that aimless revision would be useless. In the content of that theory, recall that Rawls held that nonideal theory could not be left to wander aimlessly without ideal theory to provide the goal toward which the former must work. Taking a step back, the need for the revision of that theory must also be guided by something beyond itself. At the beginning of his final work, *Justice as Fairness*, Rawls finally spells out four central roles of the discipline of political philosophy. I take it that his own theory, if it is to be acceptable, must also fulfill these roles.

The practical function

This is the role of political philosophy with which Rawls begins, and it may be tempting to see it as the discipline's primary role, since the other three are only "noted briefly" (JF 2). I have questions, which I will address shortly, about which of these roles ought to take precedence, but the practical role of political philosophy is certainly vitally important, whether or not it trumps the others.

The practical role of political philosophy arises from the reality in our world of "divisive political conflict and the need to settle the problem of order" (JF 1). According to Rawls, it ought to solve these problems by attending to the "deeply disputed questions" at issue, and finding, if possible, "some underlying basis of philosophical and moral agreement" (JF 2). To illustrate this role of political philosophy, Rawls references the debates between the Federalists and the Anti-Federalists over the ratification of the US Constitution, as well as the discussions of slavery as an institution and of the appropriate relationship of the North and South that occurred before the US Civil War.

Rawls mentions these cases without providing much discussion around them. Yet these strike me as interesting examples for Rawls to have chosen because, while the first of them seems to be an instance of political philosophy successfully filling its practical role, the second seems much more like a failed attempt. In the first case, the Anti-Federalists initially strongly opposed and refused to ratify the US Constitution which the Federalists had proposed as the governing document of the new United States of America. Instead, they supported the Articles of the Confederation, which called for a smaller Federal government and placed the bulk of political power squarely in the hands of the individual states. Despite the deep commitments of each group, however, a string of conventions and a public exchange of views (which led, roughly to both the *Federalist Papers* and *The Anti-Federalist Papers*) led to successful compromise. Both groups were ultimately willing to ratify a Constitution with an attached Bill of Rights which was meant to limit the power of the Federal Government and assure the protection of individual liberties. In the second case, however, such a common ground could not be reached. The southern Confederacy refused to consider the abolition of slavery, and wanted to separate themselves from the United States of America as it stood, while the northern Union demanded both abolition and the continuance of the Union between North and South. In this case, the conflict proved intractable, and these deep disagreements led to the US Civil War.

It is worth exploring for a moment why this second example of practical political philosophy might have failed. The divisive political conflict to which political philosophy is meant to attend lies at the intersection of real, particular, and opposed positions.

However, the search for political and moral agreement which Rawls offers as the solution to that conflict may have to be extremely vague if it is to represent a true common ground between the parties involved. Rawls himself acknowledges this. He notes that, since "of course our judgments will not coincide on all questions, and in fact many if not most social issues may still be insoluble, especially if viewed in their full complexity", we must acknowledge and support the "many simplifications of justice as fairness" (TJ 453). It may be the case that such agreement can only be reached if the full complexities of the question are left out, but it is certainly worth asking how useful these impoverished agreements will be in solving the real conflicts out of which the need to seek agreement initially rose. In the case of discussions about slavery and the union of the states, which nevertheless ended in war, it seems that whatever agreement that could have been reached would not have been enough. A conception to which both parties could have genuinely agreed would have been too vague to aid in reaching a real-world solution. This is not to say that political philosophy should not fill this practical role, nor is it to say that attempting to find common moral and philosophical ground is an inappropriate or useless endeavor. It is merely to say that finding this agreement, if it is in the case in question so vague, will not be enough. Political philosophy may well need to join forces with other disciplines in order to fulfill its practical role in these cases.

Orientation

The second role which Rawls suggests is that of orientation. In this case, political philosophy ought to help us to think of ourselves as more than individuals contingently involved in the governing of a nation. Instead, it should help us to see ourselves as members of a society with central common aims and goals. Rawls says that it is a necessity for us to see ourselves as members of a democracy with the political status of equal citizens, for recognition of this status will fundamentally change our orientation in respect to our social world. Once we see each other as members in a common project, we may be offered a "unified framework within which proposed answers to divisive questions can be made consistent and the insights gained from different kinds of cases

can be brought to bear on one another and extended to other cases" (JF 3). This strikes me as perhaps a second way in which we could understand the political philosophy which proceeded the Civil War to have failed. It seems plausible that subgroups within a society, as well as individuals, can be oriented toward that society. The member states of the Confederacy, at least, no longer saw themselves as engaged in a common project with the member states of the Union. Although this certainly does not mean that the individual citizens involved saw themselves as individuals rather than as citizens of a larger society, it seems that they constituted their society very differently than did those individuals in the Union. Perhaps then, we can see this as a failure of political philosophy only in the case of the northern aim of remaining a single union.

Reconciliation

In a rare poetic moment, Rawls tells us that the third role of political philosophy ought to be that of reconciling us to our social and political reality. In his words, it

> may try to calm our frustration and rage against our society and its history by showing us the way in which its institutions, when properly understood from a philosophical point of view, are rational, and developed over time as they did to attain their present, rational form (JF 3).

In other words, we are to appreciate and engage with our political reality, understanding how it has changed and how it could change in the future through our participation. Above all, we must "affirm our social world positively, and not merely … be resigned to it" (Ibid.). The Confederacy, then, was neither resigned to their political reality, nor willing to affirm that reality positively. Instead, they rejected their previous political reality, but recognizing that they could have a hand in forming their political future, moved to do so. While this was perhaps a failure for political philosophy in its reconciliatory role, the rebellion of the South (although perhaps not its final failure) may well be an illustration of the success of political philosophy in its final role: the realistically utopian one, to which we now turn.

Political philosophy as realistically utopian

Politically philosophy's realistically utopian role is to reconcile us to our future, much as its previous role is to reconcile us to our present. In addition to assuring us that our social reality is one which can reasonably be changed, realistically utopian philosophy is meant to give us hope that a *better* future can indeed be brought about. In Rawls's words, it "probes the limits of practical political possibility" (JF 4). If it does its job correctly, political philosophy should allow us the justified hope not only that our society can be different, but that we can imagine the particulars of a *more just* society that could indeed exist in the world. Hope for the future of our society is intimately connected with our beliefs about the world. More specifically, this hope "rests on the belief that the social world allows at least a decent political order, so that a reasonably just, although not perfect, democratic regime is possible" (Ibid.). In order to give us this hope, political philosophy must both consider the constraining conditions of the social world, and attempt to formulate the sorts of ideals and principles that such a society would try to realize.

The import of these central roles

Outside of these pages, Rawls offers little pointed discussion of these roles of political philosophy. Yet I think that there are several reasons to take their articulation here particularly seriously. They are, first of all, the final and most explicit articulation of what Rawls takes the responsibilities of political philosophy to be. Although the transition of justice as fairness from a comprehensive to a political doctrine is the most obvious example of the scope of the changes in his work, Rawls also recognizes that there are a number of other ambiguities and contradictions which call for similar revision to his theory. *Justice as Fairness* was written with the explicit intention of "presenting in one place an account of justice as fairness as (he) now (saw) it" (JF, xvi), in light of all of these necessary revisions. In other words, Rawls wrote *Justice as Fairness* in order to state his theory in a clear and unified way which took into account the positions of all of his former books and papers and presented the best, most coherent picture of his work as

a whole. Since it was the last piece which he wrote before he died, and since he took it to be a "more or less self-contained" reformulation of his theory, we have reason to treat it as his final considered judgment on the matter. The roles of political philosophy with which he begins should be no exception.

I think that these roles must be seen as noteworthy, however, in their content as well as in their context. These particular four roles provide one of the most solid groundings in practical concerns to be found anywhere in his work. All four roles of political philosophy relate crucially to society as a whole. Not a single one of them can be filled by deliberation that remains either purely theoretical or exclusively within the philosophical community. The practical role of political philosophy, after all, springs from "divisive political conflict" and requires us to respond to the "need to settle the problem" which arises from it (JF 1). The Civil War and the American Revolution are not the only examples which Rawls gives to illustrate this need. Instead, he discusses many concrete instances of the "long periods in the history of any society during which certain basic questions lead to deep and sharp conflict" (Ibid.). The selection of these concrete examples makes it clear that Rawls means for political philosophy to address the problems that we actually find in our real world and in our histories.

While the practical role of political philosophy focuses on the real-world problems which we must solve, the roles of orientation and of reconciliation are related to the experience of the real-world citizens who make up our societies. Political philosophy will find itself a failure in each of these latter capacities if it succeeds in orienting and reconciling only political philosophers to their societies. The discipline cannot meaningfully do what it is meant to in the case of orientation – cannot "contribute to how a people think of their political and social institutions as a whole, and their basic aims and purposes as a society with a history—a nation—as opposed to their aims and purposes as individuals" (JF 2) – if the majority of those who make up that people remain unaffected by the discipline's content. Nor can it do what reconciliation demands: it cannot "calm our frustration and rage against our society and its history" (JF 3) if the proper understanding of the philosophical point of view which is meant to do this remains the purview of a few academics.

Finally, there is the realistically utopian role of political philosophy. It seems to me that the relationship between this final role and the practical world is, although the most complex, also the most important. I will address some of these complexities and difficulties below, but for now I wish to point out the importance of this final role. Although Rawls does not discuss the ways in which these roles might stand in relationship to one another, I find it useful to view them in a kind of hierarchy. Realistically utopian political philosophy must allow us to hope for a future which is importantly different from and better than our present. I think that we would do well to take this hope to provide a framework in which we can see all of these other roles of political philosophy as working in the service of the realistically utopian one. In order for us to have a better future, political philosophy must succeed in its practical role – we must find a way to solve the divisive political conflict which tears societies apart. In order for us to have a better future, political philosophy must also succeed in orienting us appropriately – if we want political and social change, we must see ourselves as citizens with the political equality, status, and relationship to one another which together allow us to work in concert for that change. And finally, in order for us to have a better future, political philosophy must succeed in reconciling us to our social world – in order for political and social change to be a possibility, we cannot merely be resigned to our social world as an unavoidable and undesirable fact. Instead, we must be able to affirm it as the positive outcome of a rational process. This in turn will allow us to see our shared political future as the superior outcome of continued rational processes. I suggest that we see the four together as creating a feedback system – progress in the first three areas provide us with the hope that constitutes the fourth, while the fourth, when acquired, encourages us to work to further the other three. We should recognize each of the first three as working in the service of creating what the fourth allows us to hope for.

Realistic utopia

Let us place this fourth role of political philosophy in terms of the concepts discussed in the first chapter of this thesis. Rawlsian ideal theory, remember, is meant to

work out the principles of justice which could best govern a real society under the circumstances most hospitable to justice's correct functioning. In order to work out these principles, ideal theory must also concern itself with what these circumstances hospitable to justice's correct functioning are, and so, what a real society hospitable to justice might look like. It seems to me, then, that realistically utopian philosophy could have two possible (and not mutually exclusive) functions in respect to ideal theory: first, it could provide the content with which we describe and formulate these circumstances, principles, and societies. In other words, political philosophy could be realistically utopian in that it is used to determine the correct content of ideal theory. Second, realistically utopian political philosophy could serve as a bridging device between our own society and one which is described by ideal theory. Here, instead of providing the content for ideal theory, it would allow us to hope for the society which ideal theory describes. In this second case, it will be hard to say whether political philosophy's realistically utopian character makes it an instance of ideal or nonideal theory. It is clear, however, that this second relationship between the two concepts will leave realistically utopian political philosophy much more intimately linked to both ideal and nonideal theory than it would have been if its primary responsibility were the constitution of ideal theory.

Understanding the function of realistic utopia will be of vital importance in the attempt to understand what the task of political philosophy ought to be. As I will show later, Rawls takes the concept of realistic utopia to have served its purpose when it fills the first of these functions – that of providing the content of ideal theory. I will argue, however, that treating this function as sufficient is unacceptable. If realistic utopia remains in the realm of ideal theory, it will remain purely in the domain of philosophy. But as we have seen, none of the roles of political philosophy laid out by Rawls will be filled by philosophy which remains esoteric to the wider society. Before we can make a case, however, for either of these understandings of the function of realistically utopian philosophy, we must investigate further the concept of realistic utopia itself.

Realistic utopia is unlike other utopias

It may be that Rawls's choice of the phrase "realistic utopia" was an unfortunate one. No only in philosophy, but in the other humanities and the social sciences as well, the concept of utopia has gained an unfavorable reputation. Orwell's *1984* and Huxley's *Brave New World* are only the most prominent of many literary examples of the dystopias to which utopian aspirations seem inevitably to lead. Philosophers' doubts about the practicability of real-world utopia make us question whether utopia could be a worthwhile goal at all. And the very real and tragic consequences of political revolution in the name of utopian Marxist thought have made political scientists wary of the value of utopian thinking in politics. I say that Rawls's choice of phrase may be unfortunate because the word "utopia" brings to mind this slew of valid concerns which in the end are simply not relevant to the way in which Rawls uses the concept.

Indeed, by "utopian", Rawls means nothing more than drawing on "political (moral) ideals, principles, and concepts to specify a reasonable and just society" (JF 14). Recall from the first chapter that Rawls is not interested in explicating a perfect conception of justice as it is laid out in heaven – he is interested in clarifying principles of justice that can govern real human beings in a real world. The central objections to the concept of utopia tend to share a common concern: namely, that utopias are impossible, and that serious attempts to achieve that impossibility will lead to authoritarian dystopias. Rawls's sense of utopia, however, is tempered by its "realistic" modifier. This modifier requires that his proposals concern themselves with governing human beings as they could actually be governed, and "utopia" itself only does the work of assuring that our realistic proposals be guided by considered ideals. Nevertheless, anti-utopian criticisms are serious ones, and I here attempt to show that what I take to be the three central such criticisms do not apply to Rawls's philosophy.

One of the major arguments against utopia in the real world is that a utopia must be a blueprint of an ideal society. Since this society is in every way ideal, it cannot change for the better – and if it cannot change for the better, it should follow that in order for it to remain a utopia, it cannot change at all. According to Northrop Frye,

> considered as a final or definitive social ideal, the utopia is a static society; and most utopians have built-in safeguards against radical alterations. This feature gives it a somewhat forbidding quality to a reader not yet committed to it (Frye 1965: 31).

Of course, *most* societies have built-in safeguards to protect against radical alterations in their structure; however, most democracies, at least, also have structures for debating the merits and content of appropriate progress in that society, and for implementing these changes when they are endorsed. The problem with a utopia is that, in order for it to *be* a utopia, these sorts of changes must be precluded. This should indeed make any proponent of democracy critical of this sort of utopian thinking.

However, Rawls's "utopia" in no way necessitates this sort of changelessness. Although Rawls means different things by the word "utopia" when he uses it in different ways (the difficulties of this usage will be addressed below), the utopia which he advocates as a final social ideal is in no way a static society. His version of a realistic utopia *is* a reasonably just constitutional democracy. The modifier "reasonably" reminds us that, even when we are guided by our political moral ideals, we will still be human beings so guided, and so even our best conceivable societies will still be somewhat imperfect. Further, the fact that Rawls's utopia is a democracy shows us that his utopia cannot be changeless. Although he is centrally committed to the stability of his realistic utopia, the

> stability of a conception does not imply that the institutions and practices of the well-ordered society do not alter. In fact, such a society will presumably contain great diversity and adopt different arrangements from time to time. In this context stability means that however institutions are changed, they still remain just or approximately so, these adjustments being called for by new social circumstances (TJ 458).

We will always be living, changing creatures in a living and changing world. For us to remain concerned with utopia, then, cannot mean that we have reached a state of changeless perfection. Instead, it must mean that we remain constantly concerned with meeting the demands of justice as the world and the circumstances of our societies change around us.

A second major argument against utopia is that it would require a society unified

behind the same ideal of the good. As Karl Popper puts it, "the Utopian approach can only be saved by the Platonic belief in one absolute and unchanging ideal" (Popper 1962: 161). Rawls, however, endorses this criticism, and says that this sort of scenario is "utopian in the pejorative sense" (JF 188). As we saw in the first chapter of this thesis, Rawls takes this criticism so seriously that he fundamentally changes his entire theory of justice to account for it. Such a utopian conception would require all members of society to be united under a single comprehensive doctrine, and Rawls has recognized that this cannot be consistent with a free, democratic, well-ordered society. His switch to political rather than comprehensive liberalism, then, ensures that his realistic utopia will not fall prey to this same utopian objection.

The final central argument against utopia with which I will engage is that any utopia would require authoritarianism and violence in order to remain intact under what Rawls would call a single comprehensive doctrine. This objection recognizes that human beings will of necessity be less perfect than the utopia in which they could hypothetically live. Only a violent imposition of utopian ideals could preserve a society in which those ideals were followed. Lyman Tower Sargent tells us that "force will be necessary either because people question the desirability of the utopia or because there is disharmony between the perfect blueprint and imperfect people" (Sargent 1982: 568-9). However, Rawls once again shares this criticism of this kind of utopian society. He refers to the following as the fact of oppression: "a continued shared adherence to one comprehensive doctrine can be maintained only by the oppressive use of state power, with all its official crimes and inevitable brutality and cruelties" (JF 34). Clearly, he sees this sort of oppression as unacceptable in a democratic society. Luckily, however, the application of this criticism to Rawls's version of realistic utopia is rendered inappropriate in the same way as the last. If the shared comprehensive doctrine governing the utopia is rejected in favor of a political conception which allows all its members to privately affirm their own comprehensive doctrines, there will be no reason for the forced imposition of any comprehensive doctrine at all. Only the political conception will need to be accepted by the society's members, and this they will be willing to do because doing so allows them to pursue whichever doctrine they find most compelling in their own lives.

Rawls's use of utopia then, does not damn his conception in the way that a traditional use of utopia would. However, the use of the term does in some cases obscure his meaning. I will now turn to a positive account of just what that meaning is.

The ambiguity of realistic utopia

Rawls does not treat the concept of realistic utopia as a problematic one, but it seems to me that there are a number of severe obstacles to its clear presentation. I identify three such obstacles here. First, the only substantive treatments of the concept are found in *The Law of Peoples* and *Justice as Fairness.* Since *The Law of Peoples* concentrates on the relationship of justice as fairness to the international arena, and *Justice as Fairness* is meant to be a self-contained reformulation of his domestic theory, the differences between the concepts of realistic utopia used in a basically parallel way in both books are sometimes obscure. Moreover, Rawls engages far more extensively with the concept in *The Law of Peoples*, and the curious reader intent upon understanding realistic utopia in the domestic sphere must sometimes attempt to reapply the international treatment of the concept there in ways which Rawls may not have intended.

Second, it is difficult to apply the concept across both context *and* time, since realistic utopia makes a fairly late debut in Rawls's work. *Justice as Fairness* and *The Law of Peoples* are not only the only books in which Rawls discusses realistic utopia – they are also the last two books which he wrote. Although *Justice as Fairness* is meant to be a comprehensive reformulation of his previous work in domestic justice, it offers a far less detailed account than the ones found in either the 1971 or the 1999 editions of *A Theory of Justice*. Anyone intent upon studying realistic utopia throughout Rawls's opus, then, must reapply that concept in two ways: on the one hand, she must apply the comments in *The Law of Peoples* to the domestic context, and those in *Justice as Fairness* to the international one. On the other, she must attempt to apply the comments in both of these books to the far more extensive body of work published before the concept of realistic utopia was broached. Since fulfilling its realistically utopian role is, for Rawls, a central responsibility of political philosophy, it will surely be useful to study

A Theory of Justice and *Political Liberalism* with this role in mind. Yet, without express mention of the phrase in either of these works, undertaking this task will once again require the reader to interpret some of these earlier works in ways perhaps unintended by Rawls. Nevertheless, even if Rawls might not approve of this stretching of the idea, engaging in the exercise will be useful and necessary in order to most fruitfully understand the concept of realistic utopia across the full spectrum of his publications. Although certain extensions may be inconsistent with particular elements of Rawls's other work, I will make every effort in my own work not to extend the concept in ways which violate the broader picture and implications of Rawls's project.

Finally, understanding realistic utopia is problematic because there are at least three different ways of getting at the idea of realistic utopia: we can have realistically utopian philosophy, a realistic utopia itself, or philosophy which fulfills its realistically utopian role. For clarity's sake, I will refer to the last of these as realistically utopian philosophy*. All of these are put across in Rawls's work, although the relationship between them is never spelled out, and features of one conception do not always follow from the features of the others. All three of these understandings can also be found in both his domestic and his international treatment of realistic utopia. Let us consider these three possibilities now, since the one (or ones) on which we choose to focus will play an important part in the answer which we give to the question of the appropriate relationship between realistic utopia and ideal theory.

For a philosophical conception to be realistically utopian, Rawls states that it must meet three conditions.[16] Although some of these conditions have been gestured to in the previous section, for clarity's sake, I will discuss them systematically here. The first and second of these are conditions which it must meet in order to be realistic, the third it must meet in order to be utopian. First, it must "rely on the actual laws of nature and achieve the kind of stability those laws allow, that is, stability for the right reasons" (LP 12-13). In other words, if, given the laws of nature, it demands a sort of stability which cannot hold, it will not be realistic. Second, in order for a liberal political conception to be

[16] These conditions are essentially parallel in the domestic and international contexts. For a comparison of the two, see *The Law of Peoples* pp. 11-19.

realistic, it must have principles and precepts which are "workable and applicable to ongoing political and social arrangements" (LP 13). This means that those principles and precepts may not be too unwieldy or abstract. Unlike the principle of utility, which asks us to make almost impossible calculations in order to determine appropriate courses of action, principles which are realistic must be applicable in a useful, if not necessarily straightforward manner. Yet this focus on the realistic does not make the conception political in the wrong way. As Chris Brown says, "a realistic utopia is not a compromise between what is realistic and what is utopian; nothing has been compromised here" (Brown: 20). Rather, the utopian condition must also be met in its entirety. It is able to do this because the nature of Rawls's specific utopian requirement allows it to be complemented, rather than confined, by his conditions of realisticness. The requirement is that, in order to be utopian, a philosophical conception must employ "political (moral) ideals, principles, and concepts to specify a reasonable and just society" (JF 14). This requirement can be met in its entirety because, once again, the contrast is against pure personal utility. A realistically utopian conception of justice must be based on principles and ideals which are endorsed on the basis of their moral merits, rather than on the contingent and transitory intersections which they might have with personal interest.

When Rawls discusses a realistic utopia itself, however, the concepts which he uses are importantly different. Here, realistic utopia, rather than referring to the conditions that a conception of justice must meet, refers to a reasonably just constitutional democracy (JF 17). In this case, in order to understand realistic utopia we must consider the characteristics which such a democracy can and must have. These will include, for example, the features of a well-ordered society, the requirement of persons seen as citizens with the two moral powers, and the condition of stability for the right reasons. In this sense, a realistic utopia will have specific conditions which can be spelled out. For instance, to ensure stability for the right reasons in such a regime, Rawls requires the following things of that society: a certain fair equality of opportunity; a decent distribution of income and wealth which allows citizens to take advantage of their basic freedoms; a government which is, or ensures through its policies that some other agent will be, employer of last resort; basic health care for all citizens; and a system which

ensures the public availability of information concerning policy and which publicly finances elections (JF 50). Here, even if the realistically utopian conception of justice described in the first sense of the term leads to a realistic utopia in this second sense – that of a reasonably just constitutional democracy – the features of the latter are certainly not necessitated by the features of the former alone. The extremely particular content found in the second conception of realistic utopia can not be derived from conditions as broad as those presented in the first conception.

Finally, the remaining way of conceiving of realistic utopia is even further removed from the former two than those two are from each other. This is the role of political philosophy which was discussed above. Here, rather than specifying particular conditions for a philosophical conception on the one hand, or specific features of a type of governing system on the other, realistically utopian philosophy* describes the aim of political philosophy itself. Again, this aim is fairly consistent across Rawls's domestic and international discussions. Here, philosophy is realistically utopian* when it

> extends what are ordinarily thought of as the limits of practical political possibility. Our hope for the future rests on the belief that the possibilities of our social world allow a reasonably just constitutional democratic society living as a member of a reasonably just Society of Peoples. An essential step to being reconciled to our social world is to see that such a Society of Peoples is indeed possible (LP 124).[17]

Although there is a fairly evident tie between extending the limits of practical political possibility and the first and second understandings of realistic utopia – since the first attempts to meld political moral ideals with the actual laws of our world, and the second gives content to the idea of our desired regime – the connection in the third conception to hope, belief, and reconciliation to our political future is an interesting and distinct element. It is this element that will, I will argue, cause realistically utopian philosophy* to yield a very different answer to the question of the appropriate relationship between realistic utopia and ideal theory.

[17] For a virtually identical quotation in the domestic context, see *Justice as Fairness* pp. 4. I have chosen to use the quotation in LP, although the international context is generally less useful to my own project, because the last line of this version explicitly highlights the connection between social reconciliation, our hope for the future, and our belief in the possibility of that future.

Realistic utopia is nonetheless problematic

However, although Rawls's conception of realistic utopia does not fall prey to many of the same arguments that can be made against more traditional u- and dystopias, it seems to me that the ambiguity of the term causes the conception itself to nonetheless remain problematic. The various potential uses of the term make a straightforward understanding of the relationship between realistic utopia and ideal theory virtually impossible. Instead, each use must be analyzed, and the relationship which springs from it compared to both the context and the spirit of Rawls's wider work. In the first two ways of cashing out the concept (in the first case, philosophy which meets the conditions for realistic utopia, and in the second, a realistic utopia itself), realistic utopia can remain squarely in the realm of ideal theory. The extent to which this is problematic only becomes clear in light of the term's third use (as philosophy which fulfills its realistically utopian role), which positions realistic utopia as a bridging device between ideal theory and the world in which we live.

When Rawls uses "realistic utopia" in the first sense, to describe the characteristics of a type of philosophy, the conditions for realisticness play an important role. He is indeed concerned with a philosophy that must be applicable to real world scenarios. However, as we should remember from the discussion of ideal theory in the first chapter, Rawlsian attention to *a* real world need not signify attention to *our* real world. Although ideal theory must take into consideration a realistic account of human nature, it does so in the context of a society effectively regulated by the circumstances most hospitable to justice's correct functioning. This is clearly not our own society. As long as realistically utopian philosophy can apply to the sorts of people who would live in a realistic utopia (in the second sense in which that term can be used), it matters very little whether or not it can apply to us.[18] This use of realistic utopia, then, remains in the realm of ideal theory – it is purely a philosophers' concept. It is their job as thinkers to

[18] It will certainly matter whether we, or our children, can *become* the kinds of people who would live in a realistic utopia. Rawls holds that we can, but this, I will argue in the last chapter, is a claim which he makes far too quickly.

determine what sort of theory of justice could best govern a society of that sort. In this case we, as citizens, have no job.

In the second case of realistic utopia, when this term is used to describe a reasonably just constitutional democracy, the concept is even more clearly one that belongs in ideal theory. This use of the term describes the content of a reasonably just constitutional democracy – and it should be very clear to us, when we look at the previously mentioned conditions for stability in such a democracy, that societies of this type do not currently exist. A realistic utopia is here a goal which we might reach, and says nothing to us about the real world as it now stands, or how we could get from here to there. Describing the content of a realistic utopia in this sense is again a philosophers' project which engages itself with wider society no more than the task of describing the character of realistically utopian philosophy did.

Let me be clear: the first two senses of realistic utopia need not be illegitimate uses of the concept simply because they remain entirely in the realm of ideal theory. Political philosophers are of course philosophers; as philosophers they will legitimately need to engage in a good deal of theoretical work. Without political theorizing about the content of the society that we want, and the characteristics of the philosophy that describes that society, it will be very hard to meet any of the goals set by those roles. The problem arises when *only* ideal theory's uses of realistic utopia are employed. Recall the roles of political philosophy which we discussed at the beginning of this chapter. I argued that, if political philosophy is to fill any of those roles, it must have profound effects on the citizenry generally, and cannot remain exclusively in philosophical circles.

Rawls is right to say that we need ideals to aim toward if our work for change is to be of any use. But we must have something more than ideals if political philosophy is to fill the roles which he set for it. Ideal theory and the goals it sets certainly have a part to play in fulfilling those roles, but without something which can bridge the gap between philosophers and the rest of the world, political philosophy will remain wholly unsuccessful in completing the task which it has been set. Realistically utopian philosophy* provides just such a bridging mechanism. By emphasizing the importance of hope for what could be, it also emphasizes the fact that we are not there yet. And by

replacing the emphasis on *us*, who are so far from where we should be, it returns us to those problems of such central importance: addressing politically divisive conflict, orienting us in terms of our fellow citizens, reconciling us to our political realities, and providing us hope for the political realities of the future.

It is in this final sense that I will use the concept of realistic utopia for the remainder of my thesis. Having noted this, I will no longer use the * designation. Unless otherwise noted, references to realistic utopia will, from now on, refer to philosophy which fulfills its realistically utopian role.

Realistic utopia, realizability, and realization

If we content ourselves with the first two uses of realistic utopia, we remain in ideal theory, and we are thrown back to the feasibility debate. By accepting the conditions of realisticness as Rawls does – by making the limitations of human beings of central importance, and by demanding that our recommendation be applicable to real scenarios – we have already taken a stand in that debate. If the limitations of human beings (the "laws of nature") can constrain the requirements of justice and of a reasonably just constitutional democracy, then we have agreed that conditions of feasibility constrain what we can require of human beings. But as long as we remain in the realm of ideal theory, we fail to ask that further and more important question: the question of realization.

Determining the conditions of realizability is an acceptable stopping point for philosophers who take their task to be nothing more than providing an answer to the feasibility debate. But if we are working on Rawls's framework, and accepting the roles of political philosophy that he proposes, then we have already accepted that political philosophy must be about something more than answering theoretical debates. If we are to fill those roles of political philosophy, then we must worry as much about realization as we do about realizability. And this, in turn, will mean adding a third condition of realisticness to Rawls's realistically utopian philosophy: in addition to demanding that it is based on the laws of nature, and that it's principles be applicable to real human beings

and real human problems, it must also provide a solution toward which we in our own imperfect world can actually and successfully work. Thomas Pogge thinks that Rawls's realistic utopia already *is* "an ideal social world that is reachable from the present on a plausible path of transition" (Pogge 2007: 27). I think that he is right to emphasize the importance of this dimension of a project like Rawls's, but I think that he is also wrong to think that this is a requirement of Rawls's project as it stands.

Concentrating on ideal theory will not tell us if or how the social world which we desire is reachable from the social world which we have. And determining if and how we can reach a realistic utopia will be crucial for the third conception of realistically utopian philosophy: since the hope that it is meant to give us is so important, the reachability of a realistic utopia cannot be merely and tenuously theoretical. Given the distance between our own world and a realistic utopia, we cannot just assume that the latter as laid out in Rawls is reachable on a path of transition from our world. Instead, in order to provide us with the hope that realistically utopian philosophy is meant to give us, we should make this reachability a central condition *for* realistically utopian philosophy. This condition, in conjunction with the four roles of political philosophy, will then require us to engage in the nonideal theory that will allow us to answer the question in a way that justifies our hope. Rather than creating a theory which we have argued is realizable (as Rawls has done), we will then be involved in the actual work of realization (which both a condition like Pogge's and Rawls's roles of political philosophy should require us to do).

Why Rawls is wrong to settle for realizability

Rawls himself at one point makes a clear statement that realization need not be a requirement of political philosophy:

> While realization is, of course, not unimportant, I believe that the very possibility of such a social order can itself reconcile us to the social world. The possibility is not mere logical possibility, but one that connects with the deep tendencies and inclinations of the social world. For so long as we believe for good reasons that a self-sustaining and reasonably just political social order both at home and abroad is possible, we can reasonably hope that we or others will someday, somewhere, achieve it; and we can then do something toward this achievement (LP, 128).

However, as I have said, if political philosophy's realistically utopian role is to reconcile us as citizens to our future as a society – rather than just to provide this reconciliation to the philosophers in our midst – then theoretical assurance of this sort will not be enough. There are at least two further questions that we must ask of Rawls, in order for realistically utopian philosophy to give us the hope that it should. They are: 1) *Can* these conditions actually obtain? And 2) Do we indeed have reason to hope that these conditions *will* actually obtain?

The first of these questions may seem to be a strange one to ask, given that Rawls has already applied two conditions for realisticness to the concept of realistic utopia. One could simply point to the claim which Rawls makes: since realistically utopian philosophy must rely on "the actual laws of nature" taking "men as they are" as the citizens to be governed, only principles which can indeed govern human beings may acceptably be chosen as part of that philosophy. And if principles of justice which can govern real human beings can be formulated, then it must be possible for the parallel conditions which allow human beings to be governed by them to apply. We must remember, however, that the "deep tendencies and inclinations of the social world" that support the possibility of realistic utopia used in the second sense (as a synonym for a reasonably just constitutional democracy) are those of a well-ordered society. But it is not citizens of a well-ordered society who must be reconciled to their political future – it is us, and we in our less-than-well-ordered society have tendencies and inclinations which are as far from ideal as the less-than-just basic institutions that inspire them. Perhaps the world of Rawls' ideal theory *could* obtain in the presence of the appropriate inclinations, and perhaps those inclinations themselves *could* conceivably be our own. For the moment, however, they are not, and to assume that they could be on the basis of the sorts of tendencies and inclinations that we *would* have in a well-ordered society would be to beg the question. Given the distance between a hypothetical well-ordered society and our own, we must first ask: *can* the deep tendencies and inclinations necessary for a well-ordered society obtain?

Ferriera shares this concern. Since we can only confidently say that Rawls's

conception of justice is possible in the context of ideal theory and a well-ordered society, he thinks that belief in its possibility will be of little use to us, in our real human world. He notes that

> If justice seemed obtainable from where we are now, that fact might reconcile us to current injustice in some way; but to say that it is obtainable in an idealized, hypothetical society only vaguely similar to our own, and which we have very little idea how to reach, seems as likely to leave us unreconciled to our actual condition as to say that it is impossible. If our belief that justice is possible is limited to ideal theory, the hope that it sustains may look rather forlorn (Ferriera 2009: 100).

I think that Ferriera is right that, as it stands, the only statements which we can make with any confidence about the possibility of justice as fairness, are those which are limited to its existence in the context of a well-ordered society. But here, once again, Ferriera's interest in feasibility and my interest in realization come apart. Since realization is my concern, it is also my concern to understand how we can *make* justice "obtainable from where we are now". And if we can do this, then it seems that our hope for the possibility of meaningful justice in our own world may no longer be so forlorn.

For this reason, I think that the second question which must be answered is to what extent it is *reasonable to hope* that, as Rawls stated in the quotation above, a "just political social order both at home and abroad is possible" and that "we or others will someday, somewhere, achieve it". Again, even if we grant that our own inclinations and tendencies (or the inclinations and tendencies of those in our future societies) can conceivably be shaped into those which would support and be supported by a well-order society, it does not seem immediately obvious that they are likely to be. Indeed, since we lack the just basic institutions that would form us appropriately, without a concerted and intentional effort toward such forming, it seems highly unlikely that our own inclinations and tendencies will accidentally come to fit the bill. Since Ferriera is concerned with the feasibility of Rawls's theory of justice as it now stands, it might seem that the answer to this question should also be a depressingly negative one. Yet if we consider hope in the context of realization rather than feasibility, I think that the answer will become different. Hope is an important concept in Rawls, and Ferriera is right to point to it. Yet I will argue

in the next chapter that it can only be valuable insofar as its connection to realization is kept firmly in mind.

If Rawls's theory cannot answer these questions, or if it gives us the wrong answers, then it will require serious revision. Remember that Rawls revised his theory of justice in light of the fact that a single shared comprehensive doctrine could not stably govern a pluralistic society in which the burdens of conscience were honored. He did this because his theory as it was previously stated was not realizable – it yielded an unacceptable answer to the question of political liberalism.[19] In the case of realistic utopia, Rawls's theory will need to be altered because, as it stands, it cannot fulfill the realistically utopian role which he said that it must: that of providing a society with hope for and reconciliation to its future. Here, the problem is once again one of realizability – but in this case, proving that philosophy's realistically utopian role can be met will require engaging precisely with realization itself. Without attempts to realize the requirements of a realistic utopia, our assurance that they can be met will be based on nothing but abstract theorizing, and this surely cannot provide our wider society with the hope which we cannot do without.

In the next chapter, I will investigate what it would mean to provide satisfactory answers to these questions. However, if we are to take seriously the requirement of realization, determining the conditions for these satisfactory answers will not be enough. In the final chapter, I will suggest some ways in which political philosophers can engage themselves meaningfully in the work of realization.

[19] Recall that this question was: "How is it possible for there to exist over time a just and stable society of free and equal citizens, who remain profoundly divided by reasonable religious, philosophical, and moral doctrines?" (PL 4).

Chapter 3: Hoping for a Realistic Utopia

At the end of the last chapter, I argued that if Rawls's philosophy is to stand, it must be able to answer two questions about the possibility of the sorts of inclinations and tendencies that would be required to support a realistic utopia. In this chapter, I attempt to answer each of these questions. The answer to the second question leads us into a discussion of the value of hope. Ultimately I conclude that Rawls's philosophy cannot provide satisfactory answers to the two relevant questions, but suggest ways in which it can be altered in order to do so.

Question 1:

Can the deep tendencies and inclinations necessary for a well-ordered society obtain?

Although the outcome of this question will play a crucial role in determining whether or not Rawls's given theory of justice can stand on its own, it may simply not be a question to which we are able to provide an answer. After all, it is both a simple matter of empirical fact, and a question about the future. Since it is a matter of empirical fact, we would need to be able to induce from evidence the possibility of these tendencies and inclinations obtaining (either that these conditions have obtained somewhere in the past or do currently obtain) in order to answer with any certainty the question of whether or not they could obtain. But since it is also a question about the future, we cannot do this. It is precisely because these conditions do not obtain now and have not obtained in the past (so far as we know), that we need to ask about the possibility of their obtaining in the future. So, if we are to rely only on empirical evidence, then lacking that evidence as we do, we are left with little more than the meager comfort that it seems at least logically possible for our own tendencies and inclinations to be as they might in a well-ordered society.

Yet it seems to me that current empirical evidence of these tendencies and

inclinations is not the only way of getting at their future possibility. Certainly if they *did* obtain we would know for certain that their obtaining was possible. But the converse does not hold. Just because they have not and do not obtain does not mean that they *cannot*. Whether or not our shared social history has been a history of progress in any meaningfully value-laden sense, it has surely been a history of immense and far-reaching social change. The variety of inclinations and tendencies to which we can presently point is as wide as the corresponding variety of political and social communities across the globe which have inspired them. What we as societies take the good to be, how we think that it is acceptable to treat human beings, and what we take the responsibilities of individuals to be, are issues about which staggeringly different claims are made in various parts of the world. Even if there is, as philosophers like Kwame Anthony Appiah or Martha Nussbaum might argue, much that unites us when it comes to shared values, there is also much – especially at the level of tendencies and inclinations – which is importantly different. This seems to me to be a more than plausible reason to think that the inclinations of human beings are neither innate nor fixed.

Further, if we look into history, we can see a similarly broad variety of ways of being over the course of time within single societies. Even in its short history, the United States has proved to be an ideal example of this trend toward change and reinvention. The 1776 claim in the Declaration of Independence that "All men are created equal" was, at the time, viciously inaccurate. It was a nation of white, male, land-owner citizens. The Suffrage and Civil Rights movements have been examples of the way in which the changing tendencies and inclinations of the nation's population have been expressed in social action and later in policy. Neither is this conversation a finished one. Healthy debate still rages in the US on many social issues, including what status we ought to grant to homosexual relationships, and whether or not citizenship should be a requirement for many of the rights and privileges which we take to be so indispensable for Americans. Taking America as only one example among many, it seems clear not only that human communities can have a broad range of tendencies and inclinations, but also that these same features of a single community can experience profound revisions and reversals. From the viewpoint of the past, I suspect that we could have anticipated with no

confidence – if we could have anticipated at all! – the sort of social community that we have today. So although the tendencies and inclinations in question do not obtain now, it does not seem impossible that they might.

Although Rawls expresses great confidence in our ability to embody these tendencies and inclinations, he nevertheless seems to recognize that we cannot accurately anticipate what is to come in our human future. He knows that "there is a question about how the limits of the practicable are discerned and what the conditions of our social world in fact are" and he acknowledges that "the limits of the possible are not given by the actual" (JF 5). I think that we can take these admissions to be either damning or heartening. On the one hand, accepting them means that we cannot know with any certainty that we shall ever be any better than we are. On the other hand, however, we need not think that our condition is either fixed or inescapable. If the limits of the possible are not given by the actual, then we have the potential to be and do things that we have as of yet not even imagined.

So what should we say about the possibility of the specific tendencies and inclinations which would be required for a well-ordered society? I think that we have no choice but to leave it an open question. We cannot in an epistemically responsible way say that our desired conditions can indeed obtain in human societies, but as the sorts of creatures that we are, I do not think that we can afford to deny the possibility either. Since we do not know, we have two choices: we can either believe that our conditions can obtain, or believe that they cannot. In both cases, we could be either right or wrong. But if we refuse to believe that the conditions can obtain, it seems safe to say that they will not. After all, we are creatures who experience some important level of control over what we can do and be. Just as I am highly unlikely to act in generous ways if I believe that I am incapable of generosity, it is highly unlikely that we will ever have the appropriate tendencies and inclinations if we truly believe that we are incapable of possessing them. If we do believe, however, that we can be the sorts of people who would be formed by life in a well-ordered society, then our openness to this possibility leaves room for us to be those people in the event that our belief is correct. So while, in order to be epistemically responsible, it seems that we must leave the question open, I

think that, pragmatically, we have little choice but to act as if the answer to our original question is an affirmative one. Can the tendencies and inclinations necessary for a well-ordered society obtain? If we care at all about whether or not we do embody them, then I think that we must at least allow the (admittedly unsubstantiated) belief that we can.

Question 2:

Do we have reason to hope that the tendencies and inclinations necessary for a well-ordered society will *ever obtain?*

In answering the previous question, we just had to bite the bullet. We cannot know whether it is possible for those conditions to obtain, but insofar as this is a question about the way that we live our lives, and insofar as that question is an important one for us, we cannot simply put it out of our minds. If the appropriate inclinations and tendencies are not possible, then it is firstly the case that we cannot know, and secondly the case that there is nothing that we can do about the matter. If, on the other hand, they are possible, then we are in luck! But this luck is not something about which we can rejoice, for once again, on the basis of the evidence that we have before us, we cannot now know that they are in fact possible.

We should remember, however, our reason for asking these questions in the first place. Rawls held that facilitating reconciliation to our political future was one of our central reasons for doing political philosophy. In other words, it is important that we allow people to reasonably hope for a future society which is well-ordered and reasonably just. But why is such hoping valuable? What good does it do? The answers to these questions should, I think, inform the way in which we go about answering the question of whether or not we ought to have hope for our own inclinations and tendencies.

Although he makes reconciliation to both our political present and past a central role of political philosophy, Rawls himself nevertheless fails to give a substantive treatment of the importantly related notion of hope. In the previous section, I tried to

establish that we must believe in the possibility of realistic utopia, and Rawls as well thinks that establishing this point is vitally important. But establishing the possibility of realistic utopia does not by itself explain the link between that concept and hope. We still do not know exactly what hope is meant to do, or why it is valuable. If we turn to Rawls for an answer to these questions, I think that his clearest statement on the subject is to be found on the final page of *The Law of Peoples.* Here he says that the *belief* in the possibility of a realistic utopia "affects our attitudes toward the world as a whole… affects us before we come to actual politics, and limits or inspires how we take part in it" (LP 128) and that the hope that we or others may someday, somewhere achieve it should serve to "banish the dangers of resignation and cynicism" (Ibid.). It seems, then, that the belief in the possibility of a realistic utopia allows us the hope that it will come about. If we do not have hope that such a society is possible, then cynicism and despair will weigh us down and rob us of the will to ever treat political systems as a viable route for achieving meaningful change toward a more just society. However, in the final lines of this book, Rawls once more emphasizes possibility over hope: if realistic utopia is not possible, and "human beings 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" (Ibid.). But is it really only the lack of possibility of realistic utopia which makes life potentially not worth living? Is it logical possibility? Or likely possibility? Is the possibility alone enough to give us hope? What then does hope add? What, in the end, is the value of hoping?

Since Rawls's discussion leaves unanswered so many questions about both the nature and value of hope, I turn now instead to the philosophical literature on the subject. I concentrate primarily on Luc Bovens's account of hope both because it seems to be the one most referenced in the literature, and because it offers the most complete account that I have found of the values of hoping. But before we can ask why hoping might be valuable, I think that we must first consider the possibility that it simply is not. Maybe Rawls has got it wrong, and hope is actually valueless after all. As Bovens points out, the skeptic about hope can not unreasonably claim that, whatever the facts of our situation are, we are always better off if we do not hope. The skeptic's argument would run in the

following way: I desire something in the future which either will or will not come about. For instance, I desire that the day will become warm. If I hope for it and it does not come about, then I will be more disappointed than I would have been if I had refrained from hoping and remained more neutrally indifferent about the possibility of a change in temperature. If I hope for it and it *does* come about then I will be no better off, for when we hope we "tend to fill in the contours in the brightest colors" (Bovens, 670). Reality, even if it provides me with the outcome that I desired, will fail to map exactly onto the details of my hopes. If I had not hoped for the event, I might have been thrilled by it, but since I did I will now instead be frustrated by the minor ways in which reality does not correspond to my prior hopes. Perhaps I hoped for a warmer day, but am now frustrated that it is cloudy or that it has become too warm even for my tastes. Either way, the skeptic claims, "I would have been better off not having hoped for anything and so it is always irrational to hope for something" (Ibid.). Hoping, on this story, invariably leads to frustration and disappointment.

I think that the skeptic might be right, if we only hoped about things like the temperature, which were entirely beyond our control. After all, this is still hope in some sense. We can hope for things beyond our control in a way which meets the requirements for what Philip Pettit calls "superficial hope": this kind of hope consists of nothing more than a belief and a desire (Pettit 2004: 153). In order to hope, we must first believe that there is a possibility that a state of affairs will come about without believing that it is certain to do so. If an event is certain to occur, then we expect it rather than hope for it, and if it is impossible for that event to occur, we may wish for it, but cannot say that we hope for it. Secondly, hope requires that we desire the outcome which we consider possible. If we did not desire it, our orientation toward that outcome could just as easily be one of dread or indifference.

Interestingly, with the emphasis that Rawls places on possibility, this superficial hope seems to be all that Rawls requires to banish the dangers of resignation and cynicism. After all, he believes that realization of a just social order is irrelevant to the value of realistically utopian philosophy. He says that "the very possibility of such a (just) social order can itself reconcile us to the social world" (LP 128). In this case, then,

it seems that all that he needs should be precisely a belief and a desire. If realization does not matter, then he should be content with the kind of superficial hope that we would have for something like a change in weather – realization of justice, for him, seems to be as out of our hands as a sunny day.

Yet if this kind of hope is enough for Rawls, then I think that the skeptic might indeed be right. If frustration and disappointment are significantly harmful, which Rawls says elsewhere that they are,[20] then it might be better if realistically utopian philosophy did not cause us to hope for a more just future society. It might be better if we simply waited to see if the possibility came about. In that case, we should continue to assert that the correct tendencies and inclinations are possible, but remain entirely indifferent on the question of whether or not they will obtain.

But then, if the skeptic is right, was Rawls simply wrong to consistently mention hope in tandem with the possibility of realistic utopia? Would he have done better merely to argue for the *possibility* and leave it at that? Or would we do better to understand hope in a different way, so that hope can indeed add something to mere possibility?

Most of the philosophers writing on the subject do require more of hope than the appropriate beliefs and desires. In addition to belief and desire, each of the following philosophers requires a third constitutive condition of hope. Philip Pettit (2004) thinks that substantial (as opposed to superficial) hope requires a "cognitive resolve" to act as if your hopes were going to be fulfilled, while Darrel Moellendorf (2006) makes motivation to act in the service of the hoped-for events constitutive of meaningful hoping. Victoria McGeer (2004) holds that hope is "an essential and distinctive feature of human agency" (McGeer 2004: 100), and so connects it intimately with choices that we make and actions that we undertake. It seems to me that all three of these philosophers require their respective third conditions because they recognize that hope, when it plays its most

[20] For Rawls, self-respect is a crucial primary good, and requires a confidence in one's ability to fulfill one's intentions (TJ 386). If we are consistently unable to realize our projects, experiencing failure and self-doubt, we will become unable to continue with our endeavors. We will then become disinvested in our projects and fail to place proper weight on the values which they express. Since persons are defined by their two moral powers – one of which is the capacity for a sense of the good – a lack of proper investment in our values will be a serious problem for Rawls. However, it seems to me that this sort of lack of investment can come just as easily from a perceived lack of one's own agency as from repetitive experiences of one's own failure.

important roles in our lives, is not about things like the weather. Instead, our most important hopes are about the things which, as McGeer says, are tied to our human agency. Our hopes are for outcomes over which we can have some control, and hope is valuable precisely because it gives us the confidence which spurs us into the action needed to bring those outcomes about.

In the case of our first question (can the deep inclinations and tendencies necessary for a well-ordered society obtain?), I noted that the answer was a simple matter of empirical fact. If the answer was no, then we could not know that this was the case, and there was nothing that we could do about the matter. If, on the other hand, the answer is yes, then it remains the case that we cannot know that it is possible for them to obtain – but it is no longer the case that there is nothing that we can do. If it is possible for these conditions to obtain, then the second question (do we have reason to hope that they *will* obtain?) follows from the first. Here, a hope which spurs us into facilitating action will certainly be useful. In the absence of empirical evidence either that the correct tendencies and inclinations can obtain or that they cannot, hope suddenly becomes very important. Without hope, and the action it would inspire us to undertake, it is almost certain that these conditions will not obtain, for even if they are possible, they are not likely to come about on their own. With hope, however, we have at least a fighting chance, for our hope will inspire us to create an environment hospitable to the tendencies and inclinations that we desire.

If all this is correct, then it seems that Rawls is right to make realistically utopian philosophy, and the hope for the future which it inspires, such a central and crucial role of political philosophy. Yet, as I noted above, Rawls also seems content with realistically utopian philosophy which inspires only a superficial hope based on mere possibility, and which does not concern itself with the ultimate realization of our political ideals. As we should recall from the first chapter, the particulars of realization are the responsibility of nonideal theory as transition which concerns itself with the ways in which we might move from the world we have now to the world that we want. Since Rawls does not concern himself with this kind of nonideal theory, and instead remains almost entirely in the realm of ideal theory, the realistic utopia about which he speaks is a thing which we

might hope for as we would hope for a sunny day – both outcomes are entirely beyond our control. But instead of siding with the skeptic and claiming that it is better for us not to hope, I think that we should continue to investigate the ways in which hope can be valuable, and see if Rawls's account and this sort of hope can then be made consistent.

To determine, then, whether hope should indeed be important to Rawls's work, we will first need a clearer understanding of why hope is valuable. In addition to asking why hope is valuable generally, however, we must ask several other questions of hoping on Rawls's account: Can we understand Rawls to be talking about a hope which has the values of substantive hope? Or is he, as we feared, talking about a hope which is only superficial? Even if it is a superficial hope, can it nevertheless confer some of the values of a more substantive conception? And if it could not confer these values, would his work be compatible with a more substantial account which would bolster his project after all? Ultimately, how can we understand hope to do the work that Rawls needs it to do?

As I said, I will undertake the discussion of the value of hope primarily in light of Bovens's account. Bovens takes hope to have at least three kinds of instrumental value in addition to its intrinsic value. I shall begin here with the instrumental features of his account.

Our hopes have a causal relationship with the world

First, along with Pettit, Moellendorf, and McGeer, Bovens holds that there is a causal dependency between hoping for something and bringing it about. He offers the example of hoping for his daughter's help with a project. Although the skeptic would claim that he was better off not hoping for her help whether or not she ultimately decided to lend him a hand, Bovens points out that his hope will encourage him to act in ways which might influence her decision. His hope that she will work with him motivates him to ask her to do so – and a request from a loved one is often quite effective in motivating us to action. Even if we do not depend upon the goodwill of a loved one, however, our hopes can affect the way that we interact with the world. If I genuinely hope that I will be able to complete a marathon, then I will train for it more vigorously and enthusiastically

than I would if I felt either dejected about or attached no importance to my prospects of finishing. In Bovens's words, the skeptic is wrong about hope in this and so many other cases because "the states of the world (viz., whether I will or will not bring the task to a successful end) are causally dependent on my choice (viz., whether to hope or not to hope)" (Bovens 1999: 671). Here, hope is valuable because it increases our chances of having our hopes realized.

It seems to me, however, that hope as it stands in Rawls's project is not valuable in this way. In the project that Rawls undertakes (although not in the purposes which he assigns to political philosophy), his task is to formalize the conditions for a realistic utopia. He does not attempt, nor does he actively encourage other philosophers to attempt, the project of nonideal theory as transition. If we *were* to do this – if we were to propose ways in which we as philosophers could design avenues for change, and we as citizens could implement them – we would indeed have a situation in which the state of the world were (at least partly) causally dependent upon the actions which our hope inspired. This, however, would require us to move beyond questions of realizability into the work of realization. There can be no causal dependence between a state of the world and a proposal which is feasible but unapplied. Insofar, then, as we took a more just world to be valuable, our hope would also be valuable on the basis of its role in realizing that more just world. But since Rawls does not either do or encourage this, hope as it stands in his project only makes us vulnerable to the disappointment and disheartenment that the skeptic promises.

Hope aids us in more than fair gambles

Hope's second instrumental value, for Bovens, lies in its ability to help us to counteract risk aversion in the case of more than fair gambles.[21] He tells us that "the players who adopt a resolution to accept life's more than fair gambles tend to come out as winners, while the players who resist such a resolution tend to come out as losers"

[21] A fair gamble is a gamble in which there is a 50-50 chance of success. In a more than fair gamble, chances of success are higher, and in a less than fair gamble, they are lower.

(Bovens 1999: 672). They come out as losers, of course, because of the clichéd but importantly true fact that, in order to win, you need to play. This is by no means to say that hope would be valuable if it encouraged us to engage in just *any* gamble – while it is true to say that you cannot win the lottery unless you play, a "hopeful" strategy in this sort of case leads many to spend money with which they cannot afford to part on an infinitesimally small chance at a great return. Yet when we are presented with the opportunity for *more than fair* gambles, hope plays a crucial role in ensuring a good outcome. More than fair gambles remain gambles, so there of course remains the possibility of losing out, especially in the short term. But if one engages in more than fair gambles the odds are significantly in favor of one's coming out on top. If one engages in these gambles *consistently,* the long-term odds are even better. Hope then, helps us to counteract this risk aversion so that we enter the game in the first place. Of course, one might object that, given a more than fair gamble, a computer would also play the odds. Clearly hope plays no role here. This seems right to me, but it also seems right to say that a computer would not "play it safe", refusing to gamble at all on the basis of the very real possibility of losing out even in a more than fair gamble. Human beings certainly do play it safe in this way, refusing to take even quite good odds. It is here that hope plays its important role of counteracting risk aversion.

I think, however, that the sort of hope that Rawls talks about once again fails to be valuable in this way. As we noted, hope is only valuable when it helps us to counteract risk aversion in *more than fair* gambles. When our gambles are less than fair, as they are in the case of the lottery, acting on the basis of our hope can lead to ruinous consequences. Here, Bovens recognizes that fear which helps us to focus on our potential (and likely) losses in *less than fair* gambles is just as valuable to us as hope in more than fair gambles can be (Ibid.). There may be intrinsic value to the pleasure of hoping,[22] but hoping in the absence of at least some good reason seems to render such hope importantly valueless. Not only are the consequences of such foolish hope likely to be far too painful to be counteracted by the anticipatory pleasure of that hope, but hoping

[22] A possibility which I will discuss below.

against odds which one knows to be less than fair may involve self deception which has a negative value in its own right.

In Rawls's case, I am unconvinced by the idea that our gamble is more than fair. What we have is an unsubstantiated claim that it is possible for the object of our hopes – the tendencies and inclinations that support a well-ordered society – to obtain. Remember that these tendencies and inclinations, in Rawls's picture, are themselves the product of a well-ordered society. We have no well-ordered societies now, nor have we had them in the past. Why ought we to think that they will spring fully formed out of the air and provide us with citizens whose inclinations and tendencies are as we want them to be? Indeed, if one were to play devil's advocate, one might even argue that tragedies like the Holocaust and the genocides in Darfur and Rwanda give us serious reason to think that the tendencies and inclinations that we want will and can *not* obtain.

I realize that I have said two things here which may appear to be contradictory. Just now I have said that I am unconvinced that our gamble on a well-ordered society is more than fair, and also that we ought not be recommended by hope to act on less than fair gambles. Earlier, I said that I think that, despite the fact that we cannot know if our desired conditions can obtain (in other words, although we do not know if the gamble is more than fair), we ought to act on the basis of the hope that they can. I think, however, that despite appearances, these statements are compatible. Indeed, I do not know if Rawls's project alone offers us a more or less than fair gamble. I do think, however, that if we hope in the sense that we did when we claimed that our hopes had a causal effect on the world, then we tip the scales in our favor – we *create* a more than fair gamble. Of course, how much more than fair that gamble is, is debatable. But the harder that we work at stacking the odds in our favor – the harder we try to create situations which we encourage deep tendencies and inclinations like the ones that we want – the more that gamble moves from "fair" toward "certain".

Rawls's problem, then, remains the same, and this is only one more consequence of it. He makes a merely theoretical claim about the possibility of a well-ordered society, which he cannot back up with any empirical evidence. Such an uncertain claim can do nothing to increase the odds in our gamble. Attention to the real inclinations and

tendencies which people have, on the other hand, and to how these can be changed for the better, is more likely than anything else to increase these odds. Once more, though, simply discussing the realizability of these tendencies and inclinations will do nothing to bring them about. Active work in realization must be undertaken if providing ourselves with a more than fair gamble is our goal. So it is once again insofar as hoping encourages us to act in these ways, and not insofar as it is changes our beliefs about possibilities in the world, that it is valuable.

Hope engenders new constitutive hopes

Finally, Bovens holds that hope is instrumentally valuable in that it causes us to analyze our hopes, and so our values and goals and the kinds of people that we take ourselves to be (Bovens 1999: 673).[23] When we invest ourselves in active, engaged hope for a particular outcome, the attention and reflection which we give to our hopes extends to ourselves. In hoping for a promotion, I imagine as well my reactions and pleasures upon receiving it. From this I may realize that I had hoped for the promotion for the wrong reasons, and ought to pursue either more ethical or more pragmatic ways of reaching my goals. Alternatively, I may realize that this promotion is more important to me than I had ever imagined, and hope further to continue on my path in more explicit ways than I could have before conceived. Or perhaps in analyzing my hopes, I realize that I have been selfish in my attempts to attain them, and hope to become the sort of person who does not routinely place her own good ahead of that of others. Hope in this story is valuable because it brings us knowledge of ourselves and our values which previously eluded us.

Although this characteristic of hope may be extremely valuable in everyday scenarios, it seems to be no more valuable in the context of Rawls's project than were the others. In our day to day lives, we are prone to hoping without thinking deeply, and so it seems right that the energy that we spend on visualizing our triumphs and our reactions to

[23] Bovens takes this aspect of hope to be both intrinsically and instrumentally valuable. I do not address these two forms of value separately because Rawls's work fails to benefit in the same ways in regards to each.

them would be well-spent. In a theory like Rawls's, however, goals, principles, and values have been analyzed extensively. Justice as fairness has undergone an impressive evolution from its first appearance in *A Theory of Justice*, to its central reworking in *Political Liberalism*, through its international extension in *The Law of Peoples*, and finally to its last expression in *Justice as Fairness*. Each change has been the result of detailed and critical thought about each of the components involved in the theory, and about the implications and values of each of its parts. It seems highly unlikely, given the deep philosophical thought involved in Rawls's work, that simple hoping will clarify his theories or their goals much further. Indeed, this is one of the greatest strengths of Rawls's work – although he may not tell us *how* we can acceptably get to where we want to be, the clarity of his goals is impeccable.

Although we as political philosophers may have no need of hope in this capacity, one could say that hope is still valuable in this role insofar as political philosophers attempt to engender in the public the hope for a future well-ordered society. I think that this claim is importantly right – but it cannot be a suggestion in support of Rawls's theory. In the second chapter of this thesis I argued that, although all four of Rawls's roles of political philosophy demanded extension into the broader community, his work did nothing to facilitate that extension. I think that hope of this type may then be a crucial aspect of political philosophy, but I do not think that Rawls's philosophy provides it. Once again, if this role is going to be filled, political philosophy will need to turn its attention to realization in the form of intellectual engagement with the community on the question of and the need for justice.

The pleasures of hoping

In addition to the instrumental values of hope, Bovens also takes hope to be valuable in its own right on the basis of the pleasures that accompany hoping. Not only do we delight in the anticipation of a desired outcome, but we also benefit from the fact that "in times of hardship, there is welcome respite in hoping" (Bovens 1999: 676). One of the most obvious examples of the value of hope is that of the Jewish prisoners in

World War II's Nazi concentration camps. Hope is a common theme in biographical accounts of the time, and in many of them, life is made possible just insofar as hope remains. When hope begins to wane, life itself becomes literally impossible.

This may be the central reason for which Rawls takes hope to be so indispensable. Remember how important reconciliation was for Rawls: he cared deeply that our belief in the possibility of a realistic utopia prevent us from becoming bitter toward our world, or painfully resigned to it. Hope was the mechanism by which we could avoid these debilitating emotions and Rawls felt that realizability was enough to give it to us; for him, merely "seeing that the conditions of a social world at least allow for [the possibility of a well-ordered society] affects our view of the world itself and our attitude toward it" (JF 38). What hope must most importantly do is allow us to see the world as no longer "hopelessly hostile, a world in which the will to dominate and oppressive cruelties, abetted by prejudice and folly, must inevitably prevail" (Ibid.). For this, he thinks that the practical possibility of such a world alone will suffice. I agree with him wholeheartedly that such despair is deadly poison for the sorts of creatures that we human beings are – but I do not think that practical possibility alone can be enough to give us the hope which can sustain us.

Paulo Freire tells us that it is precisely from high hope that we fall to our deepest lows. For him, "hopelessness is but hope that has lost its bearings" (Freire 1992: 8), and "hopelessness and despair are both the consequence and the cause of inaction or immobilism" (Freire 1992: 9). We must be careful not to hope in an empty way, or we may "experience it in a mistaken form, and thereby allow it to slip toward hopelessness and despair" (Ibid.). Here, the most tragic despair is a consequence of realizing that one has hoped for something which one cannot, or has done nothing to, bring about. For Freire, it is impossible to successfully do what Rawls attempts to do: to tie the crucial human need for hope to a future which is not causally dependent upon our actions.

> Hope, as an ontological need, demands an anchoring in practice. As an ontological need, hope needs practice in order to become historical concreteness. That is why there is no hope in sheer hopefulness. The hoped-for is not attained by dint of raw hoping. Just to hope is to hope in vain (Ibid.).

Rawls may be right to emphasize the importance of hope for protecting us from tragic despair, but if Freire is right, an empty hope is the last thing which may give us this protection. Instead, such hope robs us not only of our optimism, but of our agency as well. And without a sense of our own agency, we would be very different sorts of beings from the human beings which we are now. I think that Freire is right to identify such a strong link between successful hope and our sense of agency. Although the pleasures of groundless hoping may initially be valuable, the cost is simply too great – we cannot afford the paralyzing despair that comes on the heels of empty hoping.

McGeer holds that hoping is a condition for the very possibility of living a human life, and for this reason she distinguishes not between hope and the failure to hope, but between hoping well and hoping badly (McGeer 2004: 102). Divorced from human agency as it is, the sort of empty hope that suffices for Rawls's purposes would count for McGeer as an instance of the latter kind. She notes as valuable the psychological definition of hope as "a cognitive activity that involves setting concrete goals, finding pathways to achieve those goals, and tapping one's willpower or agency to move along pathways to the specified goals" (McGeer2004: 103). Rawls's philosophy meets only the first of these conditions, so I take the hope that it inspires to fall into the category of what McGeer calls "wishful hoping":

> Wishful hopers thus generate hopes that are fanciful insofar as they are not grounded in any real understanding of how they will be realized; they are simply the direct output of desires and so undisciplined by knowledge of the world. Moreover, because wishful hopers have a dependence on external powers for bringing their hopes about, this generates a kind of passivity with respect to invoking their own powers of agency for realizing their hoped-for ends: wishful hopers await their future goods; they do not constructively work toward them… Thus, wishful hopers are highly vulnerable to despair. Because their hopes are unrealistic,[24] they are quite unlikely to be realized (McGeer 2004: 113).

For McGeer then, as well as for Freire, the wrong sort of hope is worse than no hope at all, because it sets us up for invariable despair when we realize that our own failure to exercise our agency has set the object of our hopes beyond our reach. Once again, the

[24] Wishful hopers are unrealistic in hoping not because their ends themselves are unrealistic, but because they are unrealistic about the way in which hopes are achieved – that is, they fail to realize that our hopes are achieved through our working to bring them about.

intrinsic value of pleasure in hoping does not seem to be enough to justify its high price.

The shortcomings of Bovens's account of hope

I think that McGeer is right to say that we ask the wrong question when we ask whether or not we should hope. We are creatures importantly oriented toward the future, and we have a deeply vested interest in what that future happens to be like. McGeer is right that insofar as we are meaningfully human, we *cannot help* but hope. But if hope opens us up to the tragic despair about which she and Freire speak, then we must indeed ask what it means to hope well. The problem with Rawls's philosophy is not that he addresses hope when he should not – it is that his model only provides us with tools for hoping badly.

Throughout this discussion of hope, I have focused on Bovens's account for the reasons that I mentioned above: it is the account most referenced in the literature, and it provides an extremely clear and detailed inventory of the values of hope. Nevertheless, I think that it has a serious flaw. Like Pettit, Moellendorf, and McGeer, Bovens thinks that a third constitutive condition is required in order to turn a superficial account of hope into a substantial one. His condition, however, is very different from any of theirs. For Pettit, Moellendorf, and McGeer, the third condition for hope is linked closely with our agency. For Bovens, this third condition seems to me to be more like a willingness to engage in daydreaming. For him, in addition to an uncertain belief in the possibility of an outcome and a desire for that outcome, one who hopes must also engage in "mental imaging": a "devotion of mental energy to what it would be like if some projected state of the world were to materialize" (Bovens 1999: 674). In other words, one cannot simply believe that a state of affairs is possible and consciously or unconsciously desire it. In order to hope, one must have invested time and thought into yearning for that outcome.

This condition, however, seems to me to be insufficient for valuable hope, for it fails to clarify the difference between substantial hoping and wishful hoping. Wishful hoping is much like wishful thinking. In the latter, "one's desire that *p* renders one biased in favor of evidence for *p*" (McLaughlin 1988: 42). Further, in wishful thinking one

adopts a desired belief even when the evidence is insufficient: wishing means that "one 'jumped' *beyond the evidence* to the conclusion that *p*, prompted by one's desire that *p*, and the desire that *p* bolstered the belief *despite the absence of adequate evidence*" (Ibid., italics are my own). In wishful hoping, one does not necessarily adopt the *belief* that the desired state of affairs will come about, but one does fervently hope far beyond the evidence, while failing to do anything to make that hope realized. And this seems to me to be practically irrational in the same way. Bovens's mental imaging condition, by engaging us intimately with our desires, feeds our willingness to believe that our hopes will be realized, but does nothing whatsoever to ensure that they will be. It opens us up to dangers of wishful hoping, without furthering our desired project in any meaningful way.

It seems to me that substantial hope, on the other hand, is not practically irrational in the same way, although it also fails to take the evidence before it exactly at face value. This relationship to evidence, however, can be made sense of if we take substantial hoping to parallel positive thinking, which is itself a species of motivated belief. Positive thinking is "thinking in which, for one or another reason, the thinker holds that the desired outcome is made more likely if he believes in it" (Johnston 1988: 74). In this case, the positive thinker recognizes the actual likelihood of her desired outcome, but chooses nevertheless to believe something subtly different because she holds that this outcome will in fact be made more likely by her belief. The classic example here is William James's, in which a person who comes to a gorge adopts the motivated belief that she can jump across it (James 1895). This belief quells her anxiety and so puts her in a more competent place to undertake a successful jump. The person who hopes in a substantial way does not similarly adopt the *belief* that her hope will be realized, but the relationship between her hope and the evidence is parallel to that of the relationship between the positive thinker's belief and its own evidence. Both the substantial hoper and the positive thinker move beyond the available evidence not simply because they want the desired state of affairs, but because they take it to be the case that adopting their respective positions actually makes it more likely that their desired states of affairs will come about. In this case, then, unlike the last, hoping is indeed practically rational.

Bovens's third condition for hope is not useful because mental imaging seems to

me to be sufficient only for the sort of hope that parallels wishful thinking. Yet only hope that parallels positive thinking allows us to act, practically rationally, on the basis of something which does not exactly correlate to the evidence which we have before us. What we need in order both to differentiate between substantial hoping and wishful hoping, and to preserve what is valuable about the former, is a condition more like Pettit's, Moellendorf's, or McGeer's. This sort of condition ties hope to our agency and our effective motivation to act in the service of the hoped-for outcome. Indeed, I think that even on Bovens's own account, this further condition is required. None of the benefits which Bovens takes hope to have remain if our hopes are unconnected to our actions. There can be no causal dependency between hoping and bringing about a state of affairs if hope is not tied concretely to action. Similarly, hope will not aid us in counteracting risk aversion if it does not effectively motivate us to take the more than fair gamble before us. Neither will the pleasures of hope remain sweet if our hopes are consistently unmet, and for them to be met, more often than not we will need to engage actively in bringing them about. It may be the case that hope can still be valuable in bringing about self-knowledge, but it is not clear that wishing cannot do this equally.

What makes hope distinctive and valuable, then, requires a condition connected to agency even on Bovens's own account. He merely fails to provide it. And without this sort of third condition, it is correct that we cannot make a useful link between hope and the possibility of justice in a world quite far from our own. It is precisely because Ferriera relies on Bovens's conception of hope, and does not demand a conception of hope meaningfully tied to agency, that he is correct in his estimation that our hope will be so forlorn. To give ourselves a more robust conception of hope, we must tie our hope to a condition which demands an orientation toward realization. I prefer Moellendorf's formulation of such a condition, which he refers to as the "practical aspect of hope". It states that "the taking of one's beliefs and desires as reasons for action is a necessary condition of hoping" (Moellendorf 2006: 420). It recognizes that as long as we fail to be motivated by our hopes, we cannot properly be said to be hoping at all. As long as we base our hope on the feasibility of our desired outcome, and fail to be motivated to attempt that outcome's realization, our hope will be a forlorn exercise indeed. So long as

hope *is* tied to agency, however, hoping seems to be anything but forlorn. Instead, in this case it seems to be the most practically rational position to adopt.

Hope and epistemic irrationality

Yet even if hope is practically rational, it may nonetheless be problematic. Quite simply, even if hope is practically rational in all of the ways that have been discussed, does it not still fail to be *epistemically* rational? Bovens thinks that wishful thinking, at least, is indeed epistemically irrational: "the wishful thinker raises the subjective probability of desired states of the world beyond what is warranted by the available evidence and lowers the subjective probability of undesirable states of the world below what is warranted by the available evidence" (Bovens 1999: 678). Is hoping the same? Is it not tantamount to allowing oneself to believe, against the evidence, that the chances of a hoped-for occurrence are higher than they in fact are?

I think that hope need not commit us to this sort of epistemic irrationality. This is because, in the case of hope, it seems to me that we should see epistemic rationality as intimately linked to the previous discussion of practical rationality. Pettit holds that it is possible to separate our hopes from our beliefs (Pettit 2004: 162). Although in hoping, I may choose to act as if the chances of my desired prospect are higher than they are, this does not mean that I *believe* them to be. Admittedly, this sounds peculiar. We generally take it that rationality requires us to act on our beliefs, not on something which contradicts them. But I think that there is a fruitful way of understanding epistemically rational hope as tied to, rather than separated from, belief. I think that we must only have a better understanding of substantial hope as akin to positive thinking in order to see how it can be epistemically rational. Instead of understanding hope as separate from our beliefs, I think that we should treat hope as a kind of willingness to act that is influenced by two things: both the evidence before us, *and* a justified belief. This is the belief that hope, not through a change in external circumstances, but because of the way in which it bolsters our resolve, can itself alter the probability of an event coming to pass. In this case, it is practically rational to *choose* to hope – we may act in a more positive way than

the situation seems to warrant, because our hope *actually changes that situation*. The new evidence before us does indeed justify our hope. Epistemic rationality in the case of hope, then, requires us to bear centrally in mind the practical benefits of hoping. I agree with Pettit that people can rationally act on hope if they are "prepared to admit to themselves and others that for very good, pragmatic reasons… they are refusing to expose themselves to the low or unstable tide of evidence" (Ibid.). Hope is neither epistemically irrational nor self-deceptive because it allows us the justified belief that hope actually changes the evidence when it comes to reason for action. But this can only be done if one is prepared to realize the pragmatic value of the link between hope and agency. In the context of a hope like Rawls's, which does not demand the same link, it seems to me that hoping *would* be epistemically irrational. Here, there is no increased sense of agency or bolstered sense of resolve which can change the situation in the way that is required for hope to be rational. Without the link between hope and agency, hope dissolves precisely into wishful thinking.

Some final words on hope and realistic utopia

Of course, our assumption thus far has been that it at least may be possible for the tendencies and inclinations that we would need to support a well-ordered society to obtain. We must also admit the sad possibility that they can not. In that case, the skeptic might be right once again – it would be better for us not to hope after all, especially given the despair that comes from hoping where hope is not warranted. Yet even granting that this may be the case, I think for two reasons that the project of hope in political philosophy is one which we should undertake.

First, the potential gains of our project far outweigh the costs. What we stand to gain is far more than the pleasures of hoping. Rather, what we stand to gain is a well-ordered society itself. Although we run the risk of despair, I think that the idea that our own world can be no better than it is, is itself a cause for the deepest kind of despair. That would be a world in which it indeed made sense to ask whether or not we should hope, rather than how to hope well. Even if, in the end, we fail, the potential gains of the project

of hoping for and working to build a well-ordered society are worth the risk. We have everything to gain.

Second, even if it is true that a well-ordered society (and the tendencies and inclinations that support it) cannot obtain, this does not mean that the same is true of something much closer to them than what we have now. We need not achieve our ideal itself in order to make our progress toward that ideal worthwhile. If we eliminate a few more forms of injustice, learn to treat each other with a little more respect and compassion, then our efforts will not have been in vain. And I take it as given that, even if we are not justified in hoping for the world of ideal theory, we are nevertheless justified in hoping for this much.

Let us return to the two possible relationships, discussed in Chapter Two, which can exist between realistically utopian philosophy and ideal theory. I said that realistically utopian philosophy could either provide the content for ideal theory, or it could work as a bridging device between our own world and the world that ideal theory describes. It seems to me then that we have a choice here, and that this choice is based entirely on the value that we place on hope. If we choose the first option – to allow realistically utopian philosophy to merely provide the content of ideal theory – then we do not value hope. In this case, we create no link between the world of ideal theory which we discuss and our own agency. We do not take the belief that it can come about and the desire that we have for it to obtain as motivation for any action on our parts. In short, we do not concern ourselves with realization. We wait, political philosophers engaged in wishful hoping, for something outside of ourselves to change the world around us. Here, we will have little choice but to reject the roles which Rawls assigns to political philosophy.

If, on the other hand, we choose the second option – to use realistically utopian philosophy as a bridging device between the world that we have and the world that we want – then we recognize the crucial value of hope to both our philosophy and our lives as human beings. We are only justified in having the substantial hope that Rawls wants realistically utopian philosophy to give us – and so only justified in being reconciled to our political future – if we take our desire for and our belief in the possibility of a well-

ordered society to be motivation to act in what ways we can to bring that society about. And a project like this will allow us to wholeheartedly embrace the roles which Rawls places at the center of political philosophy. Indeed, I think that it can *only* be undertaken if we accept that these roles are crucial to our endeavor.

In the end, I do not think that such a form of hope is incompatible with Rawls's project as it stands. It is simply not justified by it. I quote, once again, the passage from *The Law of Peoples* which discusses realization:

> While realization is, of course, not unimportant, I believe that the very possibility of such a social order can itself reconcile us to the social world. The possibility is not mere logical possibility, but one that connects with the deep tendencies and inclinations of the social world. For so long as we believe for good reasons that a self-sustaining and reasonably just political social order both at home and abroad is possible, we can reasonably hope that we or others will someday, somewhere, achieve it; and we can then do something toward this achievement (LP 128).

In the end, almost in passing, he suggests that it is at least an implication of realizability that we can indeed work toward realization. Rawls never says that we *cannot* or *should not* work for realization – he just doesn't require us to do so. And this is where he is wrong. Rawls is right that providing ourselves with hope for our political future is of the utmost importance, and that political philosophy could hardly have a more important role to play. But because he remains in ideal theory, he does not provide us with the tools that we need to create in ourselves a hope that is justified. We do not need to replace Rawls's account. Instead, we need to *complete* it. This will certainly require more of us than would a project that remains entirely in ideal theory. In this case, we will need to write a whole new sort of political philosophy: we will need to begin to write the nonideal theory as transition to which Rawls made his few references. And beyond that, we will need to engage with members of other disciplines to put into practice the recommendations that we make in our nonideal theory. We will need, at the most basic level, to reconceive the project of political philosophy. Realization must now be at its very core.

Conclusion: What Comes Next?

At this point one might want to point out that this thesis has been, in many ways, precisely what I have argued that political philosophy should not content itself with being: it has been a philosopher's project, written for other philosophers. I must admit that this is the case. I have not engaged in the work of realization in this project. But in my own defense, I take this to be a task for an entire discipline. And if the discipline is going to adopt this new task, then I think that there may well need to be philosophy written entirely for philosophers which can motivate this adoption. Perhaps this thesis should be thought of as an exercise in a kind of ideal theory: instead of providing content, it stipulates goals toward which the discipline of political philosophy should move. But unlike Rawls, I do not think that it is possible to end here. This project will be useless unless the content for moving toward those goals is provided. And I take this to be not only a necessary second component of my own task, but also the task of any political philosopher who endorses the roles that Rawls has said that political philosophy ought to fill. This further project is indeed a separate one – but it is not optional. So although I do not undertake this second task here, in order to acknowledge how important it is, I will end my thesis with a few suggestions on how it may be undertaken.

Writing nonideal theory as transition

First and foremost, political philosophers need to do more of the work of nonideal theory as transition. The work of transition from our own world to the sort of world that we want will be filled with unique moral political questions which, as I said in my introduction, political philosophers are in an exceptional position to address. Unlike the work of many politicians, the work of political philosophers will not be political in the wrong way – it will not be merely compromise between what we have and what we want. Instead, it will recognize the difficulties of justice and morality with which we must grapple, and attempt to provide novel solutions which benefit from a long history of

thought in moral and political philosophy. The questions which we must answer are many: How do we decide when to act on particular principles given by ideal theory, and when do we act for their sake?[25] When can we violate a principle in the short-term in order to promote that same principle in the long term? How do we weigh principles against one another in various concrete situations? How do we promote certain values without violating the liberty of conscience, or falling into paternalism? Where are our ideals ultimately unreachable, and when they are, what should we seek to replace them with? And these questions, of course, are just the tip of the iceberg.

One of our most crucial tasks as political philosophers will be to determine where we should begin. The deep tendencies and inclinations of human beings, I think, will be good candidates for at least two reasons. First, for Rawls's purposes to be met, the public will need to be brought into the work of political philosophy. Engaging them in the conversation over the nature of our tendencies and inclinations, and the possibility and importance of changing them for the better, seems like a good starting point. This conversation will help to do the work of orientation by connecting these features of human beings with the requirements of citizenship in a political community. At the same time, questions about our tendencies and inclinations are an accessible place to begin a philosophical conversation with nonphilosophers: they are closely tied to our everyday moral lives, and so are already of non-abstract, immediate importance to us.

Second, our tendencies and inclinations are a good place for nonideal theory as transition to begin because broad political change will only happen if we have a population that wants that change. We do not now have a population whose deepest tendencies and inclinations are toward justice, so addressing how these features can be acceptably changed will be of the utmost importance.[26] We will not be able to achieve a

[25] I borrow this useful distinction from Michael Phillips. He holds that "it is important to distinguish between acting directly on a value and acting for the sake of a value. We might characterize this distinction in relation to lists of actions that instantiate the relevant value… We act directly on the value when our act conforms to some description on this list; we act for the sake of a value when we act with the intent to create a world in which that value is more widely realized" (Phillips 1983: 565).

[26] Farid Abdel-Nour advocates undertaking a project with a similar emphasis, arguing that in the international context, at least, Rawls has failed to account for the inclinations and tendencies which people actually have. He says, "When in TJ and PL Rawls undertook the similar project of clarifying the existing overlapping consensus in liberal societies, he could yield egalitarian principles that were anchored in

just society without a population that wants that sort of society to exist. But *do* our tendencies and inclinations in fact change? And how can this happen? These are of course separate questions. As I said in Chapter Three, our inclinations and tendencies *do* change: over time, we have become inclined to view a much broader cross-section of questions of inequality as being also questions of injustice. Slavery, the position of women, increasingly questions of the status of homosexual relationships – all of these are issues on which our societal position has changed dramatically over time.

But what of the second question? How can this happen? Kwame Anthony Appiah says that "when it comes to change, what moves people is often not an argument from a principle, not a long discussion about values, but just a generally acquired new way of seeing things" (Appiah 2006: 73). This is at first disheartening – if arguments are not compelling, then what tools for social change do political philosophers really have? Judith Lichtenberg makes a similarly frustrating claim about a particular instance of changing our tendencies and inclinations: "the most practically relevant reason people are not more generous toward those in need is that other people are not more generous… if we want people to give more, we must raise the general level of giving in a society" (Lichtenberg 2004: 88). This as well is hopelessly circular. Individuals do not give because we do not have a culture of giving. If we want them to give more, we must simply create a culture of giving. But to create a culture of giving, we must convince individuals to give. She recognizes that "it is hard to get from where we are to where we want to go unless some individuals change their behavior, but if I am right individuals will not be strongly motivated to do so unless others do so as well" (Lichtenberg 2004: 94). If Lichtenberg and Appiah are right about the ways in which our tendencies and inclinations change, then the content of the changes made in them seems to be governed by little more than chance.

---

commitments and convictions that were already in place. In LP, however, in order to yield egalitarian principles, Rawls had to rely on commitments that are not yet in place in the intersocietal context, but rather are still in need of being crystallized and brought to the fore" (Abdel-Nour 1999: 328). Because of this, philosophers committed to the law of peoples must do active work to encourage it: "To this end, the advocates of such a world need to be active persuaders who do best to intrude and participate in the internal debates and controversies of particular societies" (Abdel-Nour 1999: 314). While it seems to me that he is wrong to think that a similar problem does not hold in the domestic context, his emphasis on inclinations and tendencies and their effect on the possibility of political change seems to me to be importantly right.

However, Lichtenberg also offers two suggestions for how we can extricate ourselves from this impotent circularity. First, we can "rely on some free-thinking, free-acting individuals who set an example that others are inspired or otherwise motivated to follow" (Ibid.). But this suggestion, while potentially effective, does not rescue us from our dependence on chance. These individuals may or may not come forward, and if and when they do, the rest of the population may or may not be moved to follow them. Second, she suggests that we can "think about new ways to design our institutions so that some of the problems I have described here can be overcome" (Ibid.). This seems like a much more promising route to take, for institutions, while being part of the basic structure which influences people's lives in a largely unconscious manner, are nevertheless designed (at least in best-case scenarios) in thoughtful, intentional ways.

If we are to reconsider the institutions of a Rawlsian system, we must turn to the institutions of the basic structure.[27] The basic structure is a particularly relevant place to begin a conversation about changing the tendencies and inclinations of citizens because Rawls realizes the deep influence that the basic structure has on these aspects of citizens' lives. It is precisely because citizens have grown up under a just basic structure that they will possess moral characters strong enough to resist temptations to injustice (JF 185). So if our basic structure has such a profound impact on the development of our citizens, and if we have control over the design of our basic structure, we ought to design our system very carefully in order to produce the kinds of citizens that we want. This emphasis on the basic structure, then, can free us from the chicken-or-egg problem of social change. A basic structure designed in a certain way can certainly contribute to a "generally acquired new way of seeing things" (Appiah 2006: 73), and it can also create and express certain social norms which will influence the behavior of particular individuals in society, thereby solving problems like Lichtenberg's problem of giving.

Yet there may be serious problems with this sort of social engineering of belief and value. At what point does such manipulating of the basic structure constitute

[27] Remember that the basic structure of a society is the collection its "main political, social, and economic institutions, and how they fit together into one unified system of social cooperation from one generation to the next" (PL 11).

paternalism? When does it begin to interfere with the liberty of conscience? At what point does it interfere with the ability of citizens to freely conceptualize their own forms of the good? I take it that these are all valid and indeed important questions, but I do not think that they need to damn the project of designing a basic structure with value-education in mind.

We should remember that political conceptions of justice are meant to accommodate only *reasonable* comprehensive doctrines. Those comprehensive doctrines which advocate violations of others' freedoms – including the freedom of conscience – can not be protected by a politically liberal regime. This means already that certain principles (evangelicalism, inegalitarianism) will be suppressed by such a regime and others (toleration, respect, etc) will be encouraged. Rawls recognizes that "the basic institutions those (latter) principles require inevitably encourage some ways of life and discourage others, or even exclude them all together" (PL 195). In fact, political liberalism not only allows, but indeed *requires* the encouragement of some values over others. Clearly, the values encouraged cannot be ones which correspond to an arbitrary comprehensive doctrine.[28] Rather, the sorts of values that can be affirmed are the ones which make a politically liberal regime possible; these political values include civility, tolerance, reasonableness, and the sense of fairness (PL 194). I take it that they must also include a mutual recognition of the value of other individuals as citizens with the two moral powers, and the value of giving these other citizens justice.

So long as we as a community endorse political liberalism, we will also need to endorse and encourage these values in our citizens. The alternative to endorsing some set of values is to endorse no system whatsoever – but in this case we could say nothing whatsoever about justice, and the only political system available to us would be anarchy. It is an unfortunate but necessary feature of political liberalism that some comprehensive doctrines will place greater emphasis on the political virtues, and that these comprehensive doctrines will then be more strongly "encouraged" by the basic structure of that society. This is ultimately the only way in which we can have political goals to work toward. If we refuse entirely to endorse political systems, then political philosophy

[28] Recall from Chapter One that such a society would not be stable.

is out of a job, and, more importantly, human societies are left rudderless.

Implementing nonideal theory as transition

Concretely, then, how do we design a basic structure which encourages our political values and discourages those which contradict them? This must be the second task of political philosophers concerned with realization – in addition to developing novel proposals via nonideal theory as transition, we must join together with members of other disciplines to enact those proposals. Pogge recognizes that political philosophers cannot do this sort of practical work on their own. We will indeed begin in an area which is entirely philosophical, but our same project, if it is to be successful, must end in an area which is entirely nonphilosophical. In the middle, however, it must pass through an area which is genuinely interdisciplinary (Pogge 1989: 7). This is where the expert knowledge of philosophers must engage with the very different expert knowledge of economists, politicians, engineers, educators, and many others. In successful realization of justice, both expert theoretical and expert practical knowledge will be indispensable.

There are several areas, I think, where engagement with expert practical knowledge will prove most fruitful in our project to reshape our deep tendencies and inclinations. These are the fields of law and education.

Educational reform will serve as a kind of positive project. We can educate children in ways that encourage the political virtues that Rawls endorses. These virtues, after all, are the basis of the kinds of tendencies and inclinations that inspire us to give each other justice. Providing children with a broad liberal arts education will also help us to avoid charges of paternalism. If we teach children the skills to critically evaluate the political virtues, then they will not accept these virtues merely because they have learned them by rote. Instead, they will accept them because they are the sorts of virtues required to support the kind of society which could be successfully governed by the principles chosen in the original position. Martha Nussbaum (*Poetic Justice; Cultivating Humanity*) argues that such an education should also include an emphasis on the literary imagination and on a deep and broad engagement with other cultures. It is easy to translate her project

of breeding compassion and understanding in the face of difference into more Rawlsian terms: this kind of education for tolerance will help students to learn the value of freedom of conscience, and to acknowledge and respect the existence of other reasonable comprehensive doctrines in their own societies. This will be a positive project because it will help to instill in students an understanding of the sort of citizens that we want in our society.

Legal reform, on the other hand, can serve as a negative complement to the positive project of educational reform. Here, we can emphasize once again the importance of our citizens' human rights and moral powers. We can design our laws – and their penalties – to track these values, and to emphasize the unacceptability of their violation. Nussbaum is once again a prime example of a philosopher who engages in this sort of practical political philosophical work. In *Hiding from Humanity*, she argues that we should focus our legal system on questions of guilt rather than questions of shame, in order to respect and endorse the moral capabilities of our citizens. Further, she suggests that we change our system of penalties so that we punish most harshly the sorts of crimes that violate human dignity and human rights. This is why legal reform will act as a negative project. It will express our values to our society by showing that we will not stand for their violation.

If we can successfully use these two tools to orient the deep tendencies and inclinations of our fellow citizens toward justice, then we will have a population which possesses the motivation to work toward and support the other requirements that allow a well-ordered society to stably thrive.[29] We can enact legislation that ensures a certain fair equality of political, social, and economic opportunity. We will begin to support taxation schemes, social programs, and employment restrictions that facilitate the decent distribution of income and wealth which allows citizens to take advantage of their basic freedoms. We will require that the government become, or somehow ensure that another party become, employer of last resort for its population so that no citizen need fear

[29] Recall that Rawls listed these requirements on page 50 of *Justice as Fairness*. While the list will likely not prove to be exhaustive, and while we may choose to reconceptualize or even drop some of its individual requirements, it nevertheless seems a promising list with which to begin our reform toward a reasonably just constitutional democracy.

joblessness. We will provide basic healthcare for all our citizens, so that no one need fear death or debilitation from preventable causes. And we will have a population which demands public financing of elections and high levels of transparency in the functioning of the government. As in the areas of educational and legal reform, political philosophers will need to work closely with experts in other disciplines to realize these projects – but once the tendencies and inclinations of our population have been educated toward justice, the groundwork for these further programs will have already been laid.

For the tradition of political philosophy that follows in Rawls's footsteps, the terms of its own project require it to engage with the realization of its recommendations in our real world. If this does not depend upon the intuition with which I began, they nevertheless lead in the same direction. Whether we are motivated by caring deeply about injustice, or by caring deeply about the commitments which are required by our engagement with political philosophy, the result will be the same: writing about justice will mean working for injustice's abolition.

Bibliography

Abdel-Nour, F. (1999) "From Arm's Length to Intrusion: Rawls's 'Law of Peoples' and the Challenge of Stability" *The Journal of Politics,* 61 (2).

Appiah, K. A. (2006) *Cosmopolitanism: Ethics in a World of Strangers* (New York/London: W.W. Norton & Company).

Bovens, L. (1999) "The Value of Hope" *Philosophy and Phenomenological Research,* 59 (3).

Brown, C. (2002) "The construction of a 'realistic utopia': John Rawls and international political theory" *Review of International Studies,* 28 (1).

Farrelly, C. (2007) "Justice in Ideal Theory: a Refutation" *Political Studies,* 55.

Ferriera, N. (2009) *Human Rights, Justice, and Feasibility: Does 'Ought' Imply 'Can'?* Unpublished D. Phil dissertation, Oxford University.

Freire, P. (1992) *Pedagogy of Hope: Reliving* Pedagogy of the Oppressed (New York: Continuum).

Frye, N. (1965) "Varieties of Literary Utopias" in Frank E. Manuel (ed) *Utopias and Utopian Thought* (Boston: Beacon Press).

Hobbes, T. (1996) *Leviathan (*Oxford/New York: Oxford University Press).

Huxley, A. (1932) *Brave New World* (New York: Harper & Brothers).

James, W. (1895) "Is Life Worth Living?" *The International Journal of Ethics* 6 (1).

Johnston, M. (1988) "Self-Deception and the Nature of the Mind" in Brian P. McLaughlin & Amelie Oksenberg Rorty (eds) *Perspectives on Self-Deception* (Berkeley/Los Angeles/London: University of California Press).

Kukathas, C. & Pettit, P. (1990) *Rawls: A* Theory of Justice *and its Critics* (Oxford: Polity Press).

Lichtenberg, J. (2004) "Absence and the unfond heart: why people are less giving than they might be" in Deen K. Chatterjee (ed) *The Ethics of Assistance: Morality and the Distant Needy* (Cambridge: Cambridge University Press).

McGeer, V. (2004) “The Art of Good Hope” *Annals of the American Academy of Political and Social Science,* vol. 592.

McLaughlin, B. P. (1988) “Exploring the Possibility of Self-Deception in Belief” in Brian P. McLaughlin & Amelie Oksenberg Rorty (eds) *Perspectives on Self-Deception* (Berkeley/Los Angeles/London: University of California Press).

Moellendorf, D. (2006) “Hope as a Political Virtue” *Philosophical Papers,* 35 (3).

Nussbaum, M. C. (1997) *Cultivating Humanity: A Classical Defense of Reform in Liberal Education (*Cambridge/London: Harvard University Press).

Nussbaum, M. C. (2004) *Hiding from Humanity: Disgust, Shame, and the Law* (Princeton/Oxford: Princeton University Press).

Nussbaum, M. C. (1995) *Poetic Justice: The Literary Imagination and Public Life* (Boston: Beacon Press).

Obama, B. (2006) *The Audacity of Hope* (New York: Random House).

Orwell, G. (1949) *Nineteen Eighty-Four, a Novel* (London: Secker & Warburg).

Pettit, P. (2004) “Hope and its Place in Mind” *Annals of the American Academy of Political and Social Science,* vol. 592.

Phillips, M. (1985) “Reflections on the Transition from Ideal to Non-Ideal Theory” *Nous*, 19 (4).

Plato. (1966) *Republic.* I. A. Richards (ed) (Cambridge: Cambridge University Press).

Pogge, T. (2007) *John Rawls: His Life and Theory of Justice (*Oxford/New York: Oxford University Press).

Pogge, T. (1989) *Realizing Rawls* (Ithaca/London: Cornell University Press).

Popper, K. R. (1962) *The Open Society and its Enemies* (New York: Harper & Row).

Rawls, J. (2001) *Justice as Fairness: A Restatement* (Cambridge, Mass/London: The Belknap Press of Harvard University Press).

Rawls, J. (1980) “Kantian Constructivism in Moral Theory” *Journal of Philosophy* vol. 77.

Rawls, J. (1999) *The Law of Peoples* (London/Cambridge, Mass: Harvard University Press).

Rawls, J. (1993) *Political Liberalism* (New York: Columbia University Press).

Rawls, J. (1999) *A Theory of Justice* (Cambridge, Mass: The Belknap Press of Harvard University Press).

Rorty, R. (1999) *Philosophy and Social Hope* (London: Penguin).

Sargent, L. T. (1982) "Authority and Utopia: Utopianism in Political Thought" *Polity,* 14 (4).