

ABSTRACT

Civic Friendship and a Value Assumption in Rawls

Peter J. Younger, Ph.D.

Mentor: Francis Beckwith, Ph.D.

In the late twentieth century, John Rawls reinvigorated the social contract theory in political philosophy. Previous contract theories could not explain how those bound by the social contract consent to be bound. Rawls argues that we consent to the contract *hypothetically*. If we would agree, under ideal conditions for resolving questions about the basic structure of society, to a particular social contract, then we actually consent (in the relevant sense) to the contract.

With this understanding of consent in mind, Rawls argues in two stages. First, he argues to the original position – arguing that his original position thought experiment represents the ideal conditions for resolving questions about the basic structure of society. Subsequently, Rawls argues *from* the original position - parties in the original position would agree to two principles of justice which he names justice as fairness. If both arguments are sound, then all of us give our hypothetical consent to the terms of the social contract spelled out by justice as fairness. *This dissertation argues that these arguments cannot both be sound.*

I approach Rawls' work with a specific concern – in modern American society, discourse has become increasingly uncivil. This background condition gives rise to inquiry into civic friendship – how citizens might wish their anonymous fellow-citizens well and thus give rise to more amicable social conditions. Rawlsian liberalism helps adapt an Aristotelian conception of civic friendship to modern conditions of the pluralistic nation-state. Yet this conception of civic friendship has certain important limitations.

Rawls designs the original position carefully – controversial assumptions may prevent people from acknowledging it as the ideal position, undermining the argument to the original position. But the argument from the original position requires the parties to select principles of justice from among a slate of options. This selection, like all acts, requires some ascription of value by the actor. In Rawls' arguments, the parties assume that fulfilling the rational desires of persons is choiceworthy. This is inconsistent with the requirement that the original position avoid controversial assumptions. The argument to the original position and the argument from the original position cannot both be sound.

Civic Friendship and a Value Assumption in Rawls

by

Peter J. Younger, B.S., M.A., M.A.

A Dissertation

Approved by the Department of Philosophy

Michael Beaty, Ph.D., Chairperson

Submitted to the Graduate Faculty of
Baylor University in Partial Fulfillment of the
Requirements for the Degree
of
Doctor of Philosophy

Approved by the Dissertation Committee

Francis J. Beckwith, Ph.D., Chairperson

Robert M. Baird, Ph.D.

Michael Beaty, Ph.D.

Robert B. Kruschwitz, Ph.D.

Perry L. Glanzer, Ph.D.

Accepted by the Graduate School

May 2017

J. Larry Lyon, Ph.D., Dean

Copyright © 2017 by Peter J. Younger

All rights reserved

TABLE OF CONTENTS

Acknowledgments	vii
Dedication	ix
Chapter One: Introduction	1
Why Update Civic Friendship?	3
Rawls’ Position	4
Critiquing Rawls	5
Rejecting the Argument <i>to</i> the Original Position	6
Rejecting the Argument <i>from</i> the Original Position	11
Looking Forward	14
Chapter Two: Aristotle’s Civic Friendship in Modern Pluralistic Societies	17
Civic Friendship According to Cooper’s Aristotle	19
Against Virtue Civic Friendship	24
Against Advantage Civic Friendship	29
Concerns About Mutuality	32
A Pseudo-Aristotelian Model	37
The Commitment Required	43
Chapter Three: The Rawlsian Value Assumption and Metaphysical Neutrality	46
The Need for the Thin Theory	48
Choiceworthiness	51
The Deliberation of the Parties	56
The Political Conception of the Person Cannot Guide Choice	58
The Situation and Description of the Parties Cannot Guide Choice	63
Representation Cannot Guide Choice	65
Choiceworthiness from an Imperative?	71
The Alternatives Open to the Parties Cannot Guide Choice	73
What the Parties Count As Reasons Cannot Guide Choice	75
The Information Available to the Parties Cannot Guide Choice	76
Rawls’ Sense of Metaphysical	79
Chapter Four: That An Assumption Regarding Value is Not Neutral	87
Critiquing Justice as Fairness Versus Critiquing Political Liberalism	89
Political Liberalism’s Dependence on Justice as Fairness.	92
The Metaphysical Commitment of Justice as Fairness	97
The Value Assumption, Paternalism, and Justice as Fairness	100
Connecting the Two Principles to Rawlsian Paternalism and to the Value Assumption	101
Rawlsian Paternalism – Parenting for Primary Goods	104

Moral Formation – Incompatible Comprehensive Doctrines	109
Requirements of Reasonability	111
Not My Will but Thine	114
An Example of the Reasonable Rejection of the Value Assumption . . .	118
Summary	121
Chapter Five: Conclusion	126
Bibliography	130

ACKNOWLEDGMENTS

Without the contributions of a great number of people, this dissertation would never have existed. I owe debts of gratitude to all my fellows in the Baylor University Philosophy Department for their time and generous support, both personal and professional, over the years. In particular, I wish to thank Ben Guido, J. Frank Holmes, and Nate Jackson for listening ears and insightful comments. Also to Jeb and Beth Barr, Mike and Karol Hardin, Grant Kaul, Sonny Massey, Matt Reynolds, and all the members of our Fellowship Family – my deepest gratitude for caring about me as a father, husband, and citizen and for support in all areas of life.

I am also grateful to Frank Beckwith for his guidance throughout this project. Always willing to discuss the impact that ideas have on the way we live out our lives and how we care for those around us and able to point to concrete examples in history, he has brought philosophy from the ivory tower to the conversations I have with friends and neighbors. My especial gratitude for referring me back to C. S. Lewis, whose work unified this project. And to all those who read gave of their time as members of my committee – Robert Baird, Michael Beaty, Perry Glanzer, and Robert Kruschwitz – my sincere thanks for your time and service.

I've presented portions of this dissertation in various circumstances. The Baylor Philosophy Department's Graduate Student Colloquium stands out in providing insightful criticisms that greatly helped in developing this dissertation. I also received particularly valuable input at the Society of Christian Philosophers 2012 Midwestern Regional Meeting. I am thankful not only to the participants in that meeting, but to the Baylor

Graduate School and the Philosophy Department – Michael Beaty and Marilyn McKinney in specific – for supporting such fruitful travel.

I also owe the deepest gratitude to my family for their love, support, and involvement throughout the entire process. It has been a pleasure to go through both the highs and the lows with my wife Rachael. She endured me being away from home, handled far more than her share of the 3:00 AM diaper changes, and gave of her time, energy, and tears to walk through this process with me. I cannot adequately express how much these things mean to me. During this time our family has grown, with Samuel's birth, with the conception and death of Ananias, and the adoption of Reuben, Batel, Tabitha, Ebenezer, and Toni. I'm grateful to all of my children for challenging me to think deeply about how best to care for them and help them develop virtuous character, which is an integral part of this project. Lastly, I'm thankful to my parents (George and Barb), my in-laws (Bob and Lorelei), and to my brother- and sister-in-law (Chris and Emily) for their love, support and prayer. *S.D.G.*

*With love for all my children: Samuel, Ananias, Reuben, Batel, Ebenezer, Tabitha, and
Toni*

CHAPTER ONE

Introduction

I approach John Rawls' work with a specific concern – in modern American society, discourse has become increasingly uncivil. Many people appear concerned with more than intellectually defending their chosen position, but rather with enacting coercive policies forcing fellow citizens to live in accordance with that position. The ideal of tolerance for diverse ways of life has given way to a litany of lawsuits by which individuals are being pressed to act contrary to the dictates of their conscience.¹

Rawlsian liberalism greatly esteems the virtue of tolerance, suggesting it may have resources for dealing with this breakdown in civil relations. If Rawlsian liberalism can provide a basis for friendly relations between anonymous citizens, then it offers great hope for restoring civil discourse in our modern pluralistic society. To this end, I present an updated pseudo-Aristotelian account of civic friendship in which Aristotle's thick conception of the good is replaced by the thin theory of the good that the parties in Rawls' original position are said to employ. Lamentably, I conclude that my pseudo-Aristotelian account inspired by Rawlsian liberalism cannot provide the basis for universal civic friendship. Instead, I locate a value assumption in this updated theory of civic friendship, with an analogous value assumption in Rawls' theory of justice.

¹ *Burwell v. Hobby Lobby* represents an effort to force a family morally opposed to providing abortifacient medications to financially provide for the purchase of abortifacient medications for use by their employees. In another high-profile case, Melissa Klein was forced to close her cake-making business because she refused to bake a wedding cake for an event in which a lesbian couple was to be legally married – an event she could not (citing reasons of conscience) recognize as a wedding.

Specifically, those who negatively value fulfilling the unspecified rational desires of people cannot participate in civic friendship on such a model.²

I then proceed to examine the similar commitment Rawls makes to the positive value of the unspecified rational desires of people. I argue that this value assumption creates a significant problem for Rawls' overall project of political liberalism. Either the value assumption is too strong to be admitted to the original position or too weak to do the work Rawls needs it to do in the original position. In the former case, Rawls' argument that the original position is the correct position from which to answer questions of basic justice is unsound because his argument depends on avoiding controversial metaphysical assumptions. In the latter case, Rawls' argument that the parties in the original position would choose the two principles of justice as fairness is unsound. In this dissertation, I argue that the argument *to* the original position and the argument *from* the original position cannot both be sound.

Due to the extensive influence of Rawls on contemporary political philosophy, there may not be any unique criticism of Rawls' work left to be given. This introduction will survey some general types of critiques of Rawls' project and trace the debts that my critique owes to each of these general types of criticism. But first, I will briefly discuss why an understanding of civic friendship is important and why Aristotle's account needs adaptation to the modern context.

² Here and throughout, rational desires are those desires which a person sees as being part of his overall pursuit of the good life. These desires are part of fulfilling that individual's ideal of life – so while a desire to eat pizza tonight is almost certainly not related to one's overarching understanding of the good life, the desire to pursue higher education likely is. Thus, the desire for higher education is a rational desire while the desire for pizza tonight is not.

Why Update Civic Friendship?

Aristotle gives an account of civic friendship that is, in some respects, peculiar to his situation writing within the ancient polis. The relatively small size of the polis, along with a more homogeneous population, gives rise to significant uniformity of world views. Worshipping common gods and having a common understanding of virtue bypassed certain kinds of political conflict that come about in modern pluralistic nation states. The people understood the good life in very similar ways, enough so that one could not only exhort the citizens to virtue, but one could name and encourage specific virtues – liberality, temperance, etc. In modern America, the diversity of worldviews makes this more difficult, particularly when public institutions are involved. Consider liberality – the Aristotelian virtue concerning the giving of small amounts of money. Per Aristotle, this virtue requires finding the mean between the vices of giving too much (wastefulness) and of giving too little (stinginess). Today, educational standards intended by one group to inculcate liberality may be seen by a second group to corrupt children with wastefulness, while a third group believes the standards will make children stingy. A policy to “teach the controversy” regarding evolution may meet criticisms from both creationists and non-theistic evolutionists who seek to make their view the only view presented in curricula.

The pluralistic character of the modern nation state practically guarantees that the citizens no longer share a conception of the good life. Since traditionally, the purpose of government is to help citizens pursue the good life, lacking a shared understanding of that goal is a major hindrance. Civic friendship then, as envisioned by Aristotle, may lead to people doing things to help their fellow-citizens which their fellow-citizens disapprove

of. Such paternalistic action is sufficiently unpopular today as to constitute a *reductio* in the eyes of many authors. Rawls seeks to address pluralism by making the task of government to provide a way for citizens to live harmoniously (or at least without killing each other) such that all may then pursue their understandings of the good life as best possible without interfering with their fellows. His work provides resources, as I discuss in the second chapter, for a new model of civic friendship, similar to Aristotle's in form, which does not at first appear to presume a shared world view. Upon examination, however, this new model still requires citizens to share a certain part of their world view. As I proceed to examine this requirement in Rawls' work, a critique of Rawls' project of political liberalism emerges.

Rawls' Position

Rawls' overarching project, which I refer to as political liberalism, attempts to convince audiences that the basic structure of society should be governed by two principles (which he calls justice as fairness).³ These principles are the subject of a social contract – a contract that we agree to by virtue of two theses Rawls argues we can agree upon after due reflection. First, the thesis that Rawls' original position thought experiment represents the correct position from which to determine matters of basic justice. I call Rawls' defense of this thesis the argument *to* the original position. Intuitively, the right way to discuss justice is from a position where we are concerned

³ The two principles are as follows:
“(a) Each person has the same infeasible claim to a fully adequate scheme of equal basic liberties, which scheme is compatible with the same scheme of liberties for all; and
(b) 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 (the difference principle).” John Rawls, *Justice as Fairness: A Restatement* (The Belknap Press of Harvard University Press, 2001), 42-43.

with justice itself, rather than securing a better position for ourselves in the resulting society. The original position thought experiment achieves this by imposing a *veil of ignorance*, behind which negotiations take place. The veil of ignorance removes all information about the representatives' constituents back in the real world – knowledge of technology levels, abundance of resources, social status and world-view of the constituent being represented, etc. are all stripped away. In this way, the representative cannot attempt to leverage the situation of the constituent to bargain for terms unfairly advantaging the constituent. Therefore, the original position is the correct position for determining matters of basic justice.

I call Rawls' argument for the second thesis the argument *from* the original position. Here Rawls argues that, given the constraints on information imposed by the original position, all parties will agree to the terms of justice as fairness. The two principles do the best job of guaranteeing a basic ability for each constituent to pursue the good life according to her or his own understanding. Thus, the representatives support the two principles. We actually agree (or will upon accepting Rawls' arguments) that 1) if we were in the original position we would agree to the two principles of justice as fairness and 2) that the original position is the right place to discuss basic justice. Therefore, we agree in the relevant sense to a social contract giving the two principles of justice as fairness as the fair terms of social cooperation.

Critiquing Rawls

Rawls' argument comes in two steps, which I have called the argument *to* the original position and the argument *from* the original position. This suggests two basic types of criticisms depending on which of the two arguments the critic considers

unsound. This dissertation argues that the argument *to* the original position and the argument *from* the original position cannot both be sound arguments. The value assumption (that fulfilling the desires of rational agents has positive value) is either too strong or too weak. If it is too strong, the argument *to* the original position is unsound because the original position contains controversial metaphysical assumptions. If it is too weak, then the argument *from* the original position is unsound because the parties cannot select the two principles of justice as fairness. The incompatibility could be resolved by rejecting many different premises of Rawls' arguments – however rejecting any of the premises leaves Rawls' project incomplete. Many authors have critiqued either the argument *to* or the argument *from* the original position. As such, this dissertation owes debts to many prior and more specific criticisms. I wish to acknowledge those debts and discuss them here so as to clarify how this work relates to those prior criticisms.

Rejecting the Argument to the Original Position

One controversial premise of Rawls' argument *to* the original position is that questions regarding basic justice should be answered prior to questions about the good life. Commonly referred to as the priority of the right, this position puts questions about how to interact with other people prior to questions about the nature of mankind, the purposes of mankind, or what might be the (objectively) best life for people. Questions about, for example, the existence of God and his will for his creatures (or the nature of his creatures) can and must wait until after establishing a political system to govern how people interact. Perhaps most importantly, the answers to questions about justice cannot depend on the answers given to questions about the good life – justice is freestanding and

does not rest upon the good. Several critiques of Rawls' work focus on rejecting the priority of the right over the good and this dissertation could be extended to join them.

Rawls requires the priority of the right over the good for the argument *to* the original position. If it is not the case that matters of justice can be settled without an understanding of the good, then the veil of ignorance is inappropriate. The parties would then need more, rather than less, knowledge. Rawls acknowledges that the parties must make a contrastive choice between systems of justice in the original position and thus require some understanding of terms like good and better. Rawls introduces the thin theory of the good to fill this gap.

Rawls defends the use of the thin theory of the good at length in "The Priority of the Right and Ideas of the Good", arguing that his use of the thin theory of the good is compatible with the priority of the right.⁴ This piece includes responses to critics such as Charles Larmore who argues in *Patterns of Moral Complexity* that the Kantian ideal of a person (and thus the thin theory of the good which Rawls appeals to because of this ideal) may be understood as part of a comprehensive doctrine.⁵ Charles Taylor, in *Sources of the Self*, similarly argues that the thin theory of the good may be understood as part of a comprehensive doctrine.⁶ Rawls defends such criticisms by introducing a distinction between a political conception of the good, that is, a limited conception of the good used only for limited purposes in establishing a political system, and a comprehensive

⁴ John Rawls, "The Priority of the Right and Ideas of the Good," in *Collected Papers*, ed. Samuel Freeman (Cambridge, MA: Harvard University Press, 1999), p449-472.

⁵ Charles E. Larmore, *Patterns of Moral Complexity*, Cambridge University Press. 1987. p118-130, esp. p120.

⁶ Charles Taylor, *Sources of the Self: The Making of the Modern Identity*. Harvard University Press, 1989. p89.

conception of the good. Rawls acknowledges the necessity of some limited conception of the good for political purposes, but maintains that political purposes do not require the more robust comprehensive conception of the good. The thin theory of the good, per Rawls, does not go beyond a political conception yet it is sufficient to identify an index of primary goods that citizens will need, regardless of their world view, to pursue their conception of the good. I argue in Chapter Three that some part of the information provided to the parties in the original position must include an assumption of value; without such an assumption, the parties cannot be motivated to act in any way and the argument *from* the original position fails. Nevertheless in Chapter Four I argue that an assumption strong enough to motivate the parties to action would be more than a political conception and would compromise the argument *to* the original position.

My argument can be understood as related to Taylor and Larmore's criticisms that the assumptions of the original position are too strong. If the assumptions of the original position are too strong, then the original position is not the correct position from which to answer questions of basic justice, contra Rawls. Rather than pressing the claim that the thin theory (or the conception of persons) is too strong for the argument *to* the original position, however, I argue that the thin theory cannot be *both* strong enough for the argument *from* the original position and (said univocally) weak enough for the argument *to* the original position. Rawls responds to Taylor and Larmore by arguing that the conceptions involved are merely political conceptions. But if Rawls makes these conceptions weak enough that the argument *to* the original position is sound, then he has made them too weak for the argument *from* the original position to be sound.

Michael Sandel and others also reject the priority of the right – and thus the argument *to* the original position. Sandel rejects the Rawlsian idea of an “unencumbered” self – the self which Rawls would have enter into the original position – without any essential relationships. It is essential to our experience, Sandel argues, that we participate in relationships, at least some of which arise *prior* to any political connection to a larger society. The political conception of the person, which Rawls would employ in the original position, assumes that persons choose their ends. Against this, Sandel maintains that persons are in part *constituted* by their ends – being a husband and father may be part of who I am. To speak of myself as adopting ends relating to fatherhood seems ludicrous from Sandel’s perspective – I cannot *choose* the ends of fatherhood, the ends of fatherhood are a *constitutive part* of my self.⁷ Culture, community and other pre-political relationships might provide individuals with purposes other than those chosen for themselves as autonomous persons. The parties in the original position though are stripped of these pre-political relationships – making them not persons, but something impoverished and lesser.

Critiques along similar lines can be found in Alasdair MacIntyre who objects, “In Rawls’s case, the only constraints are those that a prudent rationality would impose. Individuals are thus in both [Rawls’ and Nozick’s] accounts primary and society secondary, and the identification of individual interests is prior to, and independent of,

⁷ Michael J. Sandel, “The Procedural Republic and the Unencumbered Self” in *Political Theory: An International Journal of Political Philosophy*, Vol. 12, No. 1 (Feb.,1984), pp. 81-96. p85-87.

the construction of any moral or social bonds between them.”⁸ Michael Walzer offers a similar argument focusing on the connection of these selves with constitutive ends.⁹

In his later works, Rawls responds to these sorts of critiques by claiming that the conception of a person as unencumbered by pre-existing ends is only a political conception. We do not need to believe that persons are actually unencumbered, Kantian selves. But instead, this is the *part* of the idea of persons that is shared widely enough for inclusion in the original position. Rawls would have us believe that we may or may not have constitutive ends but, what we can all agree on for political purposes, is that we all adopt certain ends. Against this, my argument presses on the difference between an aim, as something a person chooses to pursue, and an end, which a person has by virtue of being a person, not through some exercise of her volition. Even if the political conception of the person does not depend on a contentious view of the person, it remains contentious whether the aims we choose to pursue are relevant to questions of justice, rather than ends that we might have by being what we are. In Chapter Three, I point out that Rawls takes us to agree that we all see ourselves as self-authenticating sources of valid claims – that is, as persons in his political sense. But this is insufficient for Rawls’ reasoning to continue forward because he requires the stronger claim that we *are* self-authenticating sources of valid claims. Again I do not argue that the political conception of the person is too strong. Rather, I argue that it cannot both be weak enough for the argument *to* the original position to be sound and yet strong enough for the argument *from* the original position to be sound.

⁸ Alasdair MacIntyre, *After Virtue*. 1st ed. University of Notre Dame Press, 1981. p232-233.

⁹ Michael Walzer, “The Communitarian Critique of Liberalism,” *Political Theory: An International Journal of Political Philosophy*, Vol. 18, No. 1 (Feb. 1990) p6-23.

Rejecting the Argument from the Original Position

The argument *from* the original position is susceptible to many different critiques due to the great detail Rawls gives of the reasoning of the parties in the original position in *Theory*. He cannot place justice as fairness side-by-side with every possible candidate for principles of justice. Instead, he simplifies by arguing that justice as fairness would be chosen over utilitarian principles of justice, with occasional expositions on how particular features of justice as fairness may be superior to any alternatives.¹⁰ Yet several philosophers have objected that different principles of justice should arise *from* the original position because of a flaw in Rawls' exposition of the parties' reasoning.

Amartya Sen argues that Rawls mistakenly tracks only agency (the ability to pursue one's conception of the good) and not the broader well-being of those in society – Rawls' conception of what the parties in the original position take into consideration is too narrow.¹¹ The argument *from* the original position focuses on the expectation of the worst-off in society having sufficient resources to pursue their conceptions of the good life. Per Rawls, the parties in the original position choose principles to maximize the expectation of the worst off. To do so, the parties must be able to understand the phrase “worst off” and have some means of identifying them. Rawls addresses this need by forming a list of primary goods. These goods are the all-purpose means required for people to pursue their individual conceptions of the good life, including wealth and income, rights and liberties, access to public office, and the social bases of self-respect.

¹⁰ John Rawls, *Justice As Fairness: A Restatement*. Cambridge, MA: The Belknap Press of Harvard University Press, 2003. p83.

¹¹ Amartya Sen, “Well-Being, Agency, and Freedom: The Dewey Lectures 1984,” *The Journal of Philosophy*, Vol. 82, No. 4 (Apr., 1985), pp. 169-221. p186-187.

Should the parties be more concerned about the well-being of their constituents, rather than their index of primary goods, the parties may arrive at different principles of justice. Rawls regards the simplifying assumptions involved in relying on the index of necessary goods as necessary to avoid an insoluble problem.¹²

My critique holds that, if there are no value assumptions in the original position, then the parties cannot agree to anything. Rawls interprets a result without agreement as agreement on ethical egoism – each person may do what she desires in pursuit of their understanding of the good life unconstrained by justice. Therefore, my position entails the conclusion that in the original position (without value assumptions) the parties reach a very different endpoint. Even if the parties in the original position considered the well-being of their constituents, rather than their index of primary goods, the resulting theory still faces the same dilemma I argue Rawls faces. If the assumptions of the original position are strong enough to motivate the parties to choose a theory of justice, then those assumptions are too strong for the original position to be the correct position for addressing questions of basic justice. If wellbeing is understood in a robust sense that makes the wellbeing of constituents worthy of choice, then wellbeing is too strong for a merely political conception.

Another strain of objection to the agreement Rawls argues the parties would reach is that the two principles of justice do not cover everything necessary for a society. Sibyl Schwarzenbach argues in “Rawls and Ownership: The Forgotten Category of Reproductive Labor” that the deliberations of the parties in the original position cover only productive labor, that is, labor aimed at economic productivity. Such productivity is

¹² John Rawls, *Political Liberalism*. Expanded Edition. New York: Columbia University Press, 2005. p181n8.

only one part of what a society requires to endure over time however. Reproductive labor is also required, viz. labor which aims at producing mature, thriving citizens, developing in them those qualities needed for living a flourishing life. When care is taken to attend to the sort of labor involved in taking care of others, Schwarzenbach argues that a very different agreement would be reached in the original position.¹³ My criticism of Rawls owes a debt to this feminist critique.

While Rawls specifically addresses cases of converting adults, he does not discuss the inculcation of moral values into children. Rawls focuses on what happens when one attempts to replace someone else's existing comprehensive doctrine with a different comprehensive doctrine. But in the case of raising children, no *replacement* is occurring. Parents don't pursue only a specific set of economic, emotional, or other primary goods to enable their child to pursue her desires, they also pursue a specific sort of desires that they hope the child will strive to fulfill. Children naturally acquire background beliefs and initial values from those who parent them. Furthermore, those who parent may reasonably see this formation as part and parcel of the parenting relationship. These facts stand in tension with Rawls' assumption that the parties are mutually disinterested. For the argument *to* the original position to be sound, the assumptions of the original position must be very weak. Rawls argues that the original position should be chosen precisely because it does not include contentious metaphysical assumptions. But if the assumptions of the original position are weak enough for the argument *to* the original position to remain sound, then the parties cannot be motivated to act, let alone act to affirm justice as fairness. Thus, if the assumptions are sufficiently weak, my argument

¹³ Sibyl Schwarzenbach, "Rawls and Ownership: The Forgotten Category of Reproductive Labor", *Canadian Journal of Philosophy*, Vol. 17, Supp. 1, 1987. p139-167, p139-140.

entails that the parties in the original position reach a very different outcome than Rawls argues that they would reach.

Looking Forward

In what follows I begin with an investigation of civic friendship. In chapter Two, I survey several arguments for the conclusion that civic friendship (understood according to Aristotle and read as a type of perfect or virtue friendship) should not be employed in modern, pluralistic societies. I then offer a further argument that Aristotelian civic friendship cannot function in modern society, whether understood as a species of virtue friendship or of advantage friendship. I then amend the account of civic friendship. Where Aristotle employs a robust conception of the good, I replace it with Rawls' thin theory of the good. While the resulting model can help extend the bond of civic friendship (though perhaps in an attenuated form) across a pluralistic society, there are still some citizens who will not be able to participate. The next step in my project is to examine the value assumption which prevents universal civic friendship as it appears in Rawls' own work.

Chapter Three, therefore, begins by examining what the commitment is and where it can be found. I argue that the parties in the original position must have some sort of a value ascription – they must see some state of affairs as being worthy of their choice and action to realize – in order to take the actions (deliberation and choice) that Rawls argues they will take. I examine the circumstances of the parties in the original position, both in terms of their situation and the information available to them, in an effort to find out from where they might receive such a value ascription. Rawls does not offer a possible source for this assumption that is consistent with his overall project. So I conclude that this

assumption has entered the original position unannounced by and without defense from Rawls. If the argument *from* the original position is to go forward, the parties must have this value assumption. Finally I argue that, in spite of Rawls' use of the slogan "political, not metaphysical" this assumption is a metaphysical one.

But many metaphysical assumptions are trivially so and do not seem objectionable – perhaps the assumption required by the argument *from* the original position is still weak enough that the argument *to* the original position might also be sound? Chapter four argues that the answer to this question is no. Rawls' project of political liberalism requires that all reasonable citizens be able to participate in what he calls an overlapping consensus on the two principles of justice as fairness. If some citizens cannot participate, then the argument *to* the original position is unsound because the assumptions made in the original position are too strong – the original position is thus not the right circumstance for answering questions relating to basic justice. I argue first that, in Rawls' understanding, the two principles of justice entail a particular doctrine of paternalism. So regardless of the motivation for participating in the overlapping consensus on the terms of justice as fairness, all those who participate must agree to this doctrine of paternalism.

I then argue that parents adhering to certain world views cannot accept this doctrine of paternalism because to do so would require them to raise their children in a way that cannot give full-throated expression to their understanding of the good life.¹⁴ They cannot live out their beliefs consistently before their children while still accepting

¹⁴ Rawls calls this a doctrine of paternalism. The name is inapt because his doctrine of paternalism (I will argue) fails to address parents' actions on behalf of their children, instead focusing only on non-natural interventions in which adults receive (state) paternal care due to some incapacity on their part. I have chosen throughout to keep with Rawls' usage rather than adopting a more apt term.

the Rawlsian doctrine of paternalism. Since they must reject the Rawlsian doctrine of paternalism, they must reject the principles of justice which entail that doctrine.

But Rawls does not require all people to join in an overlapping consensus – he only requires those deemed reasonable to be able to join in that consensus. For this reason, I examine Rawls' works to understand what it is for a comprehensive doctrine to be a reasonable one. The final step in my argument is to show that some comprehensive doctrines whose adherents cannot participate in the overlapping consensus are reasonable comprehensive doctrines. Therefore, if the parties in the original position are given strong enough assumptions to act as Rawls argues they will, then the argument *to* the original position is unsound because some reasonable citizens cannot participate in an overlapping consensus on the principles of justice as fairness. Principles strong enough for one are too strong for the other. But if the principles of justice are weak enough for all reasonable citizens to participate, then the parties in the original position cannot be motivated to act at all, let alone agree on the two principle of justice. Principles weak enough for the argument *to* the original position are too weak for the argument *from* the original position. In this case, the argument *from* the original position is unsound. The argument *to* the original position and the argument *from* the original position cannot both be sound.

CHAPTER TWO

Aristotle's Civic Friendship in Modern Pluralistic Societies

Civic friendship, a form of friendship between fellow-citizens who do not know each other personally, was a significant idea in Aristotle's thought about how cities uphold justice. In Aristotle's *polis* however, citizens had much more in common than they do in today's pluralistic nation states. Many thinkers argue that, because of the pluralistic nature of modern societies, friendship is no longer fit to help us understand the relationships of fellow-citizens since they can no longer be expected to share a conception of the good. In this chapter, I survey and add to these arguments, ultimately offering a pseudo-Aristotelian model for civic friendship which partly addresses the issues raised in these arguments.

I begin by giving an exposition of civic friendship according to John Cooper's reading of Aristotle.¹ He tells us that friendship is a reciprocal relationship in which each of the two parties wills the good for the other, for the sake of the other, and are aware of this mutual well-wishing. Friendship, however, comes in several flavors according to Aristotle, being based on advantage, pleasure, or virtue. While Cooper reads Aristotelian civic friendship as a form of advantage friendship, others dispute this interpretation. We

¹ I have chosen Cooper's rendering of Aristotle for several reasons. First and foremost, Cooper's thought on Aristotelian civic friendship is influenced significantly by Rawls. Additionally, Cooper's interpretation of Aristotle on friendship is highly influential, being a starting point for many current lines of inquiry in the understanding of friendship. Further, Cooper is simply a more qualified reader of Aristotle than I am, so I humbly defer to his better illumined reading.

are left with two salient options, civic friendship as advantage friendship or civic friendship as virtue friendship.

With these two possible understandings of civic friendship, I turn to arguments against civic friendship as a model for understanding our relationships with fellow-citizens. I canvas several arguments against understanding civic friendship as a virtue friendship because of the pluralistic nature of modern society. The arguments share a common feature that civic friendship as virtue friendship might license paternalistic state actions that are seen as inappropriate in a liberal and pluralistic society. It has been suggested that, for these reasons, if we are to use civic friendship to understand our civic relationships, we must understand civic friendship as being a form of advantage friendship. I take this reasoning a step further and argue that the pluralistic nature of modern societies undermines the reciprocal awareness required for Aristotelian friendship of any sort. If we cannot be reciprocally aware that our fellow-citizens will the good for us in a pluralistic society, we cannot understand our relationships with our fellow-citizens as civic friendships, even if we treat civic friendship as a species of advantage friendship.

To fill this void, I suggest a pseudo-Aristotelian alternative inspired by John Rawls' original position thought experiment. In the original position, the parties are unaware of the conception of the good their constituents accept. To guide their actions then, Rawls equips the representatives with a thin theory of the good, referred to as goodness as rationality. This theory states roughly that something is good for a person if it helps that person satisfy his just and rational desires. By replacing Aristotle's substantive notion of the good in his theory of friendship with Rawls' thin theory of the

good, I suggest a pseudo-Aristotelian understanding of friendship. I will argue that this pseudo-Aristotelian account avoids the difficulties suggested by other authors which militate against understanding our relationships with our fellow-citizens in the pluralistic state according to an Aristotelian model.

Finally, I draw attention to a limitation of the model. For an individual to engage in this pseudo-Aristotelian civic friendship relationship, the individual must have a metaphysical commitment to the value of the satisfaction of other people's rational desires. I look at the limitations that this commitment places on the use of this pseudo-Aristotelian model of civic friendship for understanding relationships between fellow-citizens. Subsequent chapters of this dissertation will then examine the place of a similar commitment to the value of satisfying people's rational desires in Rawlsian liberalism.

Civic Friendship According to Cooper's Aristotle

John Cooper puts forth an interpretation of Aristotelian civic friendship as a type of advantage friendship. According to Cooper, civic friendship holds between all the citizens of a city.² The citizenry of the city is to be understood in a broad sense, as opposed to the narrower sense in which Aristotle sometimes uses the term to refer only to those who may vote within the city. While civic friendship is friendship of a "diluted and reduced form,"³ it also "is really a *friendship*".⁴ As a species of friendship, civic

² John M. Cooper, "Political Animals and Civic Friendship." *Reason and Emotion: Essays on Ancient Moral Psychology and Ethical Theory*, 356–77. Princeton, NJ: Princeton University Press, 1999. p370.

³ John M. Cooper, "Aristotle on the Forms of Friendship." *Reason and Emotion: Essays on Ancient Moral Psychology and Ethical Theory*, 312–36. Princeton, NJ: Princeton University Press, 1999. p332.

⁴ Cooper, "Political Animals and Civic Friendship." p370.

friendship maintains several of the hallmarks of other varieties of friendship. Eleni Leontsini formalizes Cooper's account of Aristotelian personal friendship thusly,

- According to Aristotle: x and y are friends iff
- (1) x and y know each other
 - (2) x and y have mutual goodwill for the other's sake
 - (3) x and y feel affection for each other, and
 - (4) x and y recognise (2) and (3)⁵

Two of these features have specific importance to this inquiry: friends (2) will the good for one another for the sake of the other and (4) have mutual awareness of the relationship. In Cooper's words, "the friend will wish his friend whatever is good, for his own sake, and it will be mutually known to them that this well-wishing is reciprocated."⁶ Leontsini's formalization suffices for personal forms of friendship, but some effort is required to extend this formalization so as to accommodate civic friendship which lacks (1). Locating civic friendship as a species of either virtue or advantage friendship does much to shape one's reading of Aristotle's remarks about civic friendship.

Speaking of civic friendship as an instance of advantage-friendship, Cooper elaborates, "civic friendship exists where the fellow-citizens, [1] to one another's mutual knowledge, [2] like one another, that is, where each citizen wishes well (and is known to wish well) to the others, and [3] is willing to undertake to confer benefits on them, for their own sake, in consequence of recognizing that he himself is regularly benefitted by the actions of others."⁷ These three points roughly correspond with points (2)-(4) of

⁵ Eleni Leontsini, "The Motive of Society: Aristotle on Civic Friendship, Justice, and Concord." *Res Publica* 19, no. 1 (2013): 21–35. p25.

⁶ Cooper, "Aristotle on the Forms of Friendship." p321.

⁷ *Ibid.* p333.

Leontsini's formalization above, with necessary adjustments to accommodate the condition of anonymity.

In friendship, friends will the good for one another. Without a personal relationship, however, civic friends must will the good for their fellow-citizens generally (as they lack specific knowledge of the others' situations) and trust that their fellow-citizens reciprocate. Instead, the friend supports a basic structure of society, a set of political and social institutions that "provid[e] both for the material well-being and for the fullest possible development and exercise of the highest and best qualities of mind and character of the citizens."⁸ What the citizen wills for fellow-citizens does not vary based on the individual and his or her circumstances. Moreover, by willing for the institutions present in society to provide for the development and well-being of the citizens, the individual participates in the good of fellow-citizens. Cooper writes, "for where each aims in her cooperative activity at the good of the others, and not just at her own good, the good attained in the first instance by the others becomes, and is conceived of by herself as being, also a part of her own good."⁹ The citizen wills those things *presumed* to be good for the fellow-citizens, including participation in the good of the community rather than willing those things known to be good for that person on the basis of firsthand knowledge of the individual's situation.

Friends also will the good for one another for the sake of the other. When the other is not known, there is no specific person for whose sake one wills the other's good. Instead of willing the good for a person who can be referred to name, civic friendship

⁸ Cooper, "Political Animals and Civic Friendship." p375.

⁹ *Ibid.* p374.

involves willing the good for those satisfying a broad description. Consider how a woman may, upon discovering that she has a brother who was placed for adoption before her birth, wish him well for his own sake though she knows very little about him besides their common parentage. The personal characteristics of friendship get replaced with general, one size fits all characteristics, as civic friendship extends to a potentially large class of anonymous individuals, rather than extending to one known individual. Some may raise the concern that this deviates significantly from the more standard cases of friendship because the person whose good one wills is not necessarily the same person for whose sake one wills the good. One wills the good of the class of civic friends for the sake of the class of civic friends, without specifying particular members of that class in either case, because one lacks the detailed knowledge to will more specific goods for specified individuals. The set of persons for whose sake the good is willed includes each and every person whose good is being willed. For any particular individual, her good is willed, and the good is willed for her sake, insofar as they are a member of the class of fellow citizens.

Finally, the reciprocal nature of friendship differs from that of civic friendship. Because citizens do not personally know all of their fellow-citizens, they cannot know for certain that the relationship is mutual. Instead a general assumption of reciprocity becomes evident. Cooper writes, “in a community animated by civic friendship, each citizen assumes that all the others, even those hardly or not at all known to him, are willing supporters of their common institutions and willing contributors to the common social product from which he, together with all the other citizens, benefits.”¹⁰ Each

¹⁰ Cooper, “Aristotle on the Forms of Friendship.” p333.

citizen holds this assumption on the grounds that all recognize the common value of working together to support their social institutions. Each citizen partakes in the good of society as a whole and, to promote the good of society as a whole, ought then to reciprocate and wish each fellow-citizen well.¹¹ Cooper summarizes, “civic friendship, then, as the special form of friendship characteristic of this kind of community, is founded on the experience and continued expectation on the part of each citizen, of profit and advantage to himself, in common with the others, from membership in the civic association.”¹² The individual participates, at least in part, because of an assumption that others will reciprocate so that the individual will personally benefit from the arrangement of society. This assumption of reciprocity, as I shall call it, provides motivation for each citizen to maintain membership in the society. Individuals may have other reasons for participating in the civic relationship as well, but the assumption of reciprocation is important to making the civic relationship a friendship. Bearing in mind the differences between civic friendship and personal friendship, I then paraphrase Leontsini to formalize civic friendship as follows:

According to Aristotle: x and y are civic friends iff

(1*) x and y are fellow-citizens

(2*) x and y have mutual goodwill for the other’s sake

(3*) x and y feel affection for each other, and

(4*) x and y can reasonably assume (2*) and (3*)

¹¹ If this sounds like the Rawlsian good of social union, it is likely not a coincidence. Cooper notes that his interest and ideas on civic friendship were stimulated in part by conversation with Rawls. John M. Cooper, “Aristotle on the Forms of Friendship.” In *Reason and Emotion: Essays on Ancient Moral Psychology and Ethical Theory* (Princeton University Press, 1999), 332, n23.

¹² *Ibid.* p333.

Civic friendship then, according to Cooper's model, retains the basic features of personal friendship, even though citizens do not personally know their fellow-citizens. The citizen wills the good for fellow-citizens based on general assumptions about the good life and its requirements without specific knowledge of the life circumstances of the fellow-citizen. Citizens bear this good will toward their fellow-citizens as members of the class of fellow-citizens, not as individuals with whom the citizen has a relationship. Finally, the citizens assume that their fellow-citizens reciprocate out of interest in the good of society as a whole, rather than observing the reciprocal nature of the relationship. For example, consider a citizen who considers contemplation of the beautiful to be an indispensable part of the good life. This citizen may express his civic friendship for others by supporting public funding of the arts to give fellow citizens increased opportunities to contemplate the beautiful and enjoy the good life. The citizen may express this friendship by advocating for public education policy that will steer children towards proper appreciation of the arts, to help children develop the right sort of character to enjoy the good life. Though fellow-citizens do not necessarily have any personal relationship with the citizen in question, the fellow-citizens recognize his activities as promoting their good whenever they become aware of those activities and the fellow-citizens can assume that the citizen is working for their common interest and wills their good. If the fellow-citizens reciprocate this well-wishing and affection, then the citizen is their civic friend.

Against Virtue Civic Friendship

I now turn to several arguments rejecting civic friendship as a way of understanding our relationships to our fellow-citizens. Each of these arguments targets

civic friendship understood as a variety of virtue friendship. Each argument shares the theme that the pluralistic character of modern society makes Aristotelian virtue friendship an untenable model for our relationships to our fellow-citizens. Under Aristotle, friends willed *the good* for one another, for the other's sake, and with mutual knowledge of the relationship. These arguments play on the fact that, without a shared conception of the good, wishing our fellow citizens well without knowledge of their understanding of the good becomes problematic because we might act so as to bring about states which impact the other in ways he or she would not appreciate. Now I will introduce arguments against civic friendship as Aristotelian virtue friendship from Russell Bentley, David Kahane, and Sibyl Schwarzenbach.

Bentley argues that civic friendship in Aristotle should be understood as a species of advantage friendship, rather than virtue friendship. But Bentley continues, "if this argument about the differences between the forms of friendship cannot be accepted, th[is] paper argues that we should not draw on Aristotle for an understanding of civic friendship because any similarity it might have to virtue friendship would license illiberal interventions in the lives of citizens in service of some idea of moral improvement."¹³ Bentley argues that Aristotelian virtue friendship has two parts: wishing well for the other and viewing the friend as another self. The former part may account for the attractiveness of understanding civic friendship as a form of virtue friendship.¹⁴ Bentley accuses Cooper of neglecting the second part of friendship though and argues that this second part causes problems for civic friendship as any form of Aristotelian friendship

¹³ R. K. Bentley, "Civic Friendship and Thin Citizenship." *Res Publica* Vol. 19, no. 1 (2013): 5–19. p.5.

¹⁴ *Ibid.* p.13.

that involves mutual well-wishing motivated for the sake of the other.¹⁵ He argues that, when one views a fellow-citizen as another self, one cannot tolerate moral flaws in the other self and may use coercive state power to rectify character flaws perceived in the fellow-citizen. This is problematic because the alleged character flaws may not register as flaws in the eyes of the fellow-citizen. Bentley writes, “If a good life is a life lived in accordance with virtue, then, clearly, perfect Aristotelian friends will be concerned about one another’s adherence to the demands of virtue. They will watch over their friend’s conduct as strictly as they watch over their own because they want their friend to achieve the ethical ideal.”¹⁶ Seeing fellow-citizens as another self leads the members of society to take too strong an interest in the moral character of their fellow-citizens. Bentley takes issue with the substantive account of the good found in Aristotle being employed. He worries, “mutual concern for another’s good *character*, not merely another’s good, [...] is crucial to the common life of the city.”¹⁷ We cannot escape this concern for the fellow-citizen’s character though, because concern for the other’s character *is* concern for the other’s good in Aristotle. The two will not come apart because of the understanding of the good Aristotle employs. Rejecting this level of interest in fellow-citizens’ character entails rejecting civic friendship as a form of Aristotelian virtue friendship. Bentley does not reject understanding civic friendship as a form of Aristotelian advantage friendship, however, because he believes that the lesser forms of Aristotelian friendship should be understood as self-centered.

¹⁵ *Ibid.* p.13.

¹⁶ *Ibid.* p.13.

¹⁷ *Ibid.* p.9.

David Kahane offers a similar critique of civic friendship as a species of Aristotelian virtue friendship. He characterizes Aristotelian friendship as object centered, that is to say, the attractive force of friendship derives from qualities of the friend, as opposed to a capacity for empathy in the other or experiences shared with the other. This sort of friendship is premised on a sameness between the two friends – without a sameness the one cannot see the fellow-citizen as a second self or a mirror on herself. Kahane writes, “for Aristotle, then, the bonds between citizens are unequivocally premised on sameness, though properly constituted civic bonds can sustain unity given the centrifugal forces of group difference. Insofar as a city is divided by group interests it is all the more important to cultivate shared understandings of the public good, these providing the basis for civic stability and solidarity.”¹⁸ If we accept reasonable pluralism as a permanent feature of society, we ought not demand shared objective qualities of our fellow-citizens. Kahane then concludes that other understandings of friendship better correspond to civic friendship because they avoid focusing on a particular set of shared objective qualities in this way. Kahane’s argument draws a weaker conclusion than Bentley’s, concluding only that other models of friendship should be preferred over Aristotelian virtue friendship. Yet Kahane’s argument shares the premise that moral differences endanger friendship, leading to his conclusion that we ought to seek another model.¹⁹

Sibyl Schwarzenbach also argues that Aristotle’s understanding of civic friendship needs updating to model relationships in a pluralistic state. Schwarzenbach

¹⁸ David Kahane, “Diversity, Solidarity and Civic Friendship.” *Journal of Political Philosophy* 7, no. 3 (1999): 267–86. p274.

¹⁹ *Ibid.* p273.

distinguishes between the different types of Aristotelian friendship according to how concerned one friend is with the other's character.

What in fact distinguishes character friendship from the other two kinds is, rather, that the description under which one loves the other is a description of that other's whole (or near whole) character. In the case of pleasure or advantage friendships, by contrast, one loves the other friend under some particular and limited description only, that is, as someone who in general brings advantage to me or as someone who is generally fun to be around.²⁰

She holds, similarly to Bentley and Kahane, that the emphasis placed on the moral character of the fellow-citizen leaves insufficient room for the tolerance required by a pluralistic society. Schwarzenbach concludes that we must depart from Aristotelian civic friendship because, "civic friendship in the modern state – owing to the heterogeneity of the population, multiplicity of religions, and so forth – can no longer reasonably entail that citizens concern themselves with the comprehensive moral character of fellow citizens, but only now with their public political character."²¹

Schwarzenbach responds to these difficulties by attempting to create a more liberal variation of Aristotelian civic friendship that can function under conditions of pluralism – indeed the distinction between comprehensive moral character and public political character already appears to draw on Rawls. I will eventually follow Schwarzenbach in this direction, following up on her suggestions and desiderata. I do this by formalizing a pseudo-Aristotelian understanding of civic friendship that can address concerns arising from pluralistic conditions. First, however, we should investigate whether or not we really need to do that much work and whether or not we might salvage civic friendship as a form of Aristotelian advantage friendship.

²⁰ Sibyl Schwarzenbach, "On Civic Friendship." *Ethics* 107, no. 1 (October 1996): 97–128. p 100.

²¹ *Ibid.* p 113.

Against Advantage Civic Friendship

Understanding civic friendship as a species of Aristotelian advantage friendship mitigates some of the concerns raised in the previous section. As Schwarzenbach points out, advantage friendship does not depend as heavily on the comprehensive moral character of the friend, but on a limited part of the friend's character. I shall show that this move helps to address the concerns related to the pluralistic conditions that prevail in modern societies. But even if this is successful in addressing the concerns raised previously, I argue that pluralism undermines civic friendship understood as any form of Aristotelian friendship in a different way, by undermining the conditions required for mutual awareness of the friendly relationship.

I argue that Bentley misreads Cooper's understanding of civic friendship in Aristotle. Bentley objects to understanding civic friendship as a form of Aristotelian friendship of any sort that entails seeing the other as a second self on the grounds that seeing the other as a second self will make one willing to intervene in the other's life to rectify any character flaws one sees in the other. This concern overlooks a very important distinction between civic friendship and personal friendship – civic friends do not have a personal relationship. Personal friendship may give rise to paternalistic concerns for the other's character. But because the fellow-citizen remains anonymous, one cannot develop specific concerns about her comprehensive moral character. Cooper rightly minimizes what Aristotle says about the friend being another self and a mirror because, in civic friendship, the other is not seen. The fellow-citizen cannot serve as a mirror, reflecting the citizen for her own inspection because she does not see her anonymous fellow-citizen. One might receive reports of *generalities* about fellow-

citizens in other parts of the country, but cannot gain a clear enough picture of a specific fellow-citizen's character to begin attempting to address flaws specific to the other's character short of having a personal relationship with that other person.

Bentley's concern then must be based on the use of coercive government power, rather than a personal attempt to cultivate the fellow-citizen's character. Perhaps the citizen's generalized willing that his fellow-citizens would live the good life manifests itself by supporting laws and social arrangements that would favor what the citizen understands as virtue and hinder what the citizen understands as vice. After all, citizens have a concern for the character of fellow-citizens that goes beyond their concern for commercial partners.²² But Cooper goes on to give examples of what this concern for the character of fellow-citizens looks like in modern societies. Cooper writes, "the typical American when she hears, say, about the attitudes Wall-Street brokers and commercial bankers have quite routinely been holding about privileged information that comes their way in their professional work, or about sleaziness in government circles, feels injured in ways she certainly does not feel in hearing similar things said about people in high places abroad."²³ Cooper continues, "it seems that, typically, citizens even of a modern mass democracy feel tied to one another in such a way and to such an extent that they can and do take an interest in what their fellow-citizens quite generally are like as persons; they want to think of them as good, upstanding people, and definitely do not want them to be small-minded, self-absorbed, sleazy."²⁴ These modest desires

²² Cooper, "Political Animals and Civic Friendship." p366.

²³ *Ibid.* p367.

²⁴ *Ibid.* p368.

about the character of one's fellow-citizens sound more like an interest in the public political character of others, rather than a concern for their comprehensive character. When one weighs the modesty of these hopes for others and the fact that we tend to respond to these affronts with an upturned nose rather than moving for coercive legislation to ban small-mindedness or sleazy private habits, civic friendship, understood as a form of advantage friendship, does not appear objectionably paternalistic. While we might attempt to pass legislation protecting others from the practices of those Wall-Street brokers Cooper mentions, this does not appear such a great threat to the possibility of liberal society.

Treating civic friendship as a species of advantage friendship also addresses some of Kahane's concerns about difference being a threat to civic friendship. While advantage friends do have some concern for the character of the other, as discussed above, this concern does not threaten the relationship. The relationship can endure while the mutual advantage of maintaining the relationship lasts. While some ethical differences may endanger this mutual advantage, the range of such differences is narrow. In fact, this narrow range of concern moves in the direction of Schwarzenbach's proposal for a model of civic friendship that concerns itself only for the public political character of fellow-citizens, rather than caring for their comprehensive moral character. While the citizen wishes his fellow-citizens well, i.e. wishes for them to have good character, the relationship only *requires* that the fellow-citizens have good character in a very limited set of traits and does not make demands about the private lives of the fellow-citizens. Considering civic friendship as a form of advantage, rather than virtue, friendship may not eliminate all concerns raised by a pluralistic society, but mutual well-wishing does

not require that all citizens be the same as Kahane fears, nor does it require paternalistic coercion of fellow-citizens as Bentley fears.

Concerns About Mutuality

Taking the advantage friendship approach to civic friendship reduces concerns brought up by the first feature of friendship previously numbered – willing the good for the other for the other’s sake. That does not mean, however, that civic friendship, understood as a form of Aristotelian advantage friendship, can model citizens’ relations to their fellow-citizens in a pluralistic society. I argue that conditions of pluralism undermine the second condition of Aristotle’s I have called attention to – that the friends are *mutually aware* of the reciprocal well-wishing. If two citizens do not have a sufficiently similar understanding of the good, when one becomes aware of what the fellow-citizen wills for her, she will not recognize it as willing the good for her. If she does not recognize that the fellow-citizen wills the good for her, she cannot participate in mutual awareness of willing the good for one another.

In Cooper’s understanding of Aristotle, mutual awareness of the relationship is a feature of all friendships and civic friendship is definitely a friendship. Therefore, for the relationship between any two citizens to be understood as Aristotelian civic friendship, the two must mutually recognize that the other wills the good for them. This may pose an immediate problem because the two citizens do not know one another personally. In civic friendship then we should understand this willing of the good as presumed, rather than recognized, by the parties in order to account for the anonymity of the parties. Cooper explicitly states that citizens assume, rather than have direct awareness of this connection. “In a community animated by civic friendship, each citizen *assumes* that all

the others, *even those hardly or not at all known to him*, are willing supporters of their common institutions and willing contributors to the common social product, from which he, together with all the other citizens, benefits” (emphasis added).²⁵

For example, a set of parents with children in public schools may assume that other property owners within the school district support the education of the couple’s children. The couple does not have to know all the people within their school district to believe that the others wish their family well. The other property owners support the public school through their property taxes, thereby contributing to the welfare of the couples’ children. All the residents of the district contribute to providing an educational option for all children within the district and, on the basis of this contribution, the couple has evidence supporting belief in the well-wishing of anonymous fellow-citizens owning property within that school district. Until a conflict arises (say over a school voucher program which would allow public funding to pay tuition to religious schools) it is reasonable for all parties to assume mutual well-wishing.

Such an assumption can be supported by witnessing our own attitudes towards our fellow-citizens. Cooper states that civic friendship “is founded on the experience and continued expectation, on the part of each citizen, of profit and advantage to himself, in common with the others, from membership in the civic association.”²⁶ When we catch ourselves feeling more concerned about a negative event because we are told that some of our fellow-citizens were among those injured, we gather a modicum of evidence to support our belief that we mutually wish each other well. When we attempt to buy goods

²⁵ Cooper, “Aristotle on the Forms of Friendship.” p333.

²⁶ *Ibid.* p333.

made in our own country, we see that we want capital to remain with our fellow-citizens rather than going towards others with whom we lack that bond. When we see those who do not wish their fellow-citizens well, such as the Wall-Street bankers in Cooper's case above, we feel that something is wrong with such people – they have failed in some way that is not obvious if we regard them only as shrewd business people making contracts and obeying them to the letter. These pieces of evidence are very small, but there are few challenges to the idea that we, as a populace generally, wish our fellow-citizens well. In consideration of these evidences, minor though they may be, we are justified in assuming that our fellow-citizens wish us well, at least until that assumption is challenged by particular observations.

But such challenges do arise when citizens begin to learn about the diversity of comprehensive doctrines present in a pluralistic modern nation. A citizen's comprehensive doctrine lays out a set of actions to perform under the description of willing the good for a fellow-citizen. I call this set of actions the citizen's extension of well-wishing. This set of actions is generated by the conjunction of the citizen's intension of well-wishing and the citizen's comprehensive doctrine. A citizen's extension of well-wishing is indexed to the citizen, being partly determined by that citizen's comprehensive doctrine and first-person perspective, while the intension of well-wishing is not similarly indexed. For at least some cases, fellow-citizens will work in direct opposition to what the citizen perceives as his own good. If two citizens disagree in some particular area of their extensions of well-wishing, their assumption that the other wills the good for them will be called into question when they become aware of this disagreement.

Consider a society in which some citizens believe that everyone must always be provided with a guaranteed, adequate supply of nutritious food at public expense – to do less would jeopardize the well-being of citizens who could suffer from malnutrition or jeopardize their ability to *freely* accept an offer of employment. Others within this society believe that everyone must have full responsibility for feeding their family (to motivate them to having an industrious character) – to do less would jeopardize the well-being of citizens who might suffer from poor character. Those advocating for a guaranteed social minimum may fail to recognize that those in the other group are also willing what they perceive to be good for all members of society, instead taking the other group’s political stance as evidence that they do not will the good for all members of society. Likewise, those who would place restrictions on a social minimum to preserve the motivation to industry may well construe the situation as evidence that members of the other group desire a system that is systematically harmful to developing good character. Both groups may see evidence that the members of the other group do not wish their fellow-citizens well, due to a difference of what the two groups believe to be the most important values and threats to those values. This evidence may then undermine the assumption that members of the other group will the good for their fellow-citizens, thus jeopardizing the recognition of well-wishing that civic friendship requires.

This conflict between individuals’ extensions of well-wishing does not require a personal relationship to undermine mutual awareness of well-wishing. A citizen does not need to know which fellow-citizen holds some beliefs to know that the beliefs exist among his fellow-citizens. If one knows that people of religious persuasion X exist in society and knows that people of religious persuasion X will something for their fellow-

citizens, then one knows that people in society will do something for their fellow-citizens. In fact, the anonymity of those who believe may further serve to undermine confidence in mutual well-wishing – not by giving evidence to the contrary, but simply by allowing fears to flourish. If the specific individual were known, then the personal relationship would offer a context in which the citizen might be reassured that the fellow-citizen wills the good for the citizen, even if they have disparate extensions of well-wishing.

These evidences are not particularly weighty; while they may tip the metaphorical scales in support of the assumption of mutuality, little evidence is required to tip the scales against this assumption. Because of this, evidence that fellow-citizens will things for one that one holds to be incompatible with one's own good - things that contradict one's determinate conception of the good for one's own life - can quickly overwhelm the evidence supporting the assumption that our fellow citizens will the good for us. If this occurs and the individual comes to lack justification for believing that the fellow-citizen wishes him well, civic friendship based on an Aristotelian model will break down. And this sort of situation will occur in a society with a plurality of comprehensive doctrines. Nothing prevents the diversity of comprehensive doctrines from generating diverse or contradictory determinate conceptions of the good and sheer numbers of instances will practically guarantee at least some breakdowns of civic friendship. The deeper problem for civic friendship (understood as a form of Aristotelian friendship) arises not because we cannot tolerate citizens willing the good for their fellow-citizens, but because citizens cannot mutually recognize the extension of others' well-wishing for them as corresponding to the intension of well-wishing.

A Pseudo-Aristotelian Model

Civic friendship then, understood as a species of Aristotelian friendship, does not make a good model for civic relationships within a pluralistic society. Even if the theory does not guide citizens towards levels of concern for the comprehensive character of fellow-citizens that would be considered a *reductio* against such a model of civic friendship, conditions of pluralism will undermine the mutual character of Aristotelian friendship. While all citizens might will the good for their fellow-citizens, the variation in expressions of this will prevent fellow-citizens from recognizing mutual well-wishing – a necessary condition for Aristotelian friendship of any sort. Even though citizens do not have a determinate conception of the good for their anonymous fellow-citizens, citizens still aim at a comprehensive conception of the good which others in society disagree with, per the assumption of pluralism. Civic friendship on an Aristotelian model can avoid insisting upon a determinate conception of the good for fellow-citizens due to conditions of anonymity and centering the friendship on advantage rather than the character of the fellow-citizen. But for civic friendship to flourish under pluralistic conditions, we see that these relationships must also avoid aiming at a particular comprehensive good. Such a model might be built with relatively simple (though significant) changes from the Aristotelian model of civic friendship laid out by Cooper.

Cooper's model of Aristotelian civic friendship seems especially amenable to a Rawlsian liberal re-working. Cooper himself credits conversation with John Rawls as significantly informing his understanding of civic friendship.²⁷ Rawls also envisions a sort of friendship between citizens – he writes, “the acceptance of the principles of right

²⁷ *Ibid.* p332, n23.

and justice forges the bonds of civic friendship and establishes the basis of comity amidst the disparities that persist. Citizens are able to recognize one another's good faith and desire for justice even though agreement may occasionally break down on constitutional questions and most certainly on many issues of policy."²⁸ In this section, I hope to reconstruct a pseudo-Aristotelian model of civic friendship. In doing so, I leave untouched the requirements that citizens wish each other well for the sake of the others and that they are aware that this well-wishing is reciprocated by their fellow-citizens. The change comes by understanding the anonymity of the parties in the manner of Rawls' thought experiment of the original position. This new model assumes that, because the specific desires of one's fellow-citizens are unknown, one's good will for those fellow-citizens is guided in much the same way as the representatives in the original position wish well for those whom they represent. The parties in the original position as set up by Rawls employ a thin theory of the good which Rawls calls goodness as rationality. Under this theory, something is good for a person insofar as it helps that person to fulfill her rational desires. This changes the understanding of *well-wishing*. By removing Aristotle's understanding of the good and replacing it with goodness as rationality (a theory intended to avoid comprehensive commitments), the resulting theory can address the problem of mutual recognition of well-wishing.

To state this formally, I paraphrase from Leontsini again:

x and y are civic friends iff

(1**) x and y are fellow-citizens

²⁸ John Rawls, *Theory of Justice*. Reprint, Cambridge, MA: The Belknap Press of Harvard University Press, 1971. 517.

(2**) x and y both will, for the sake of their fellow-citizens, the satisfaction of their fellow-citizens' desires.

(3**) x and y feel affection for each other, and

(4**) x and y can reasonably assume (2**) and (3**)

With this adjustment, the distinction between understanding civic friendship as a form of advantage friendship or as a form of virtue friendship disappears. There is no strong conception of the good which could support something which could be called virtue friendship. Civic friendship, understood in this way, will follow more closely with Aristotle's account of advantage friendship because the account of the good employed in civic friendship is part of a political conception and not part of a (partially) comprehensive doctrine.²⁹

The move from willing a comprehensive understanding of the good to a thin understanding of the good prevents individuals from having a determinate conception of the good of an anonymous fellow-citizen. According to goodness as rationality, the things that qualify as good for the individual depend upon the desires that the individual wants to pursue. Circumstances or objects that one citizen may consider good for her determinate life plan will not necessarily appear as goods to her fellow-citizens, as those circumstances or objects may not help the fellow-citizen to fulfill his rational desires. Only a relatively short list of things will help any citizen to satisfy the desires she

²⁹ Political conceptions of the good are, in Rawlsian terms, freestanding, meaning that they are not dependent on some (partially) comprehensive doctrine. While (partially) comprehensive doctrines may attempt to lay out an understanding of the world and certain features of it, a political conception is more limited in scope, establishing usefulness for facilitating social cooperation, rather than correspondence with reality. Rawls writes, "To use a current phrase, the political conception is a module, an essential constituent part, that fits into and can be supported by various reasonable comprehensive doctrines that endure in the society regulated by it." John Rawls, *Political Liberalism* (Columbia University Press, 2005), 12n13.

happens to have. One can recognize these all-purpose means to achieving the satisfaction of rational desires as Rawls' primary goods. Just as the representatives in Rawls' original position thought experiment will a larger minimum share of primary goods for their unknown constituents, civic friends will larger shares of primary goods for their fellow-citizens, because primary goods are the only surefire way to help their fellow-citizens satisfy their rational desires.

Without specific knowledge of fellow-citizens' specific understandings of the good, citizens cannot will more specific goods for their civic friends. This does not interfere, however, with willing the good according to a comprehensive doctrine for personal friends. But when one contemplates the good of a fellow-citizen, qua civic friend, one contemplates that person's good without an understanding of the fellow-citizen's determinate conception of the good. Without the determinate conception of the good, this model of civic friendship addresses Bentley's concerns of paternalistic action. Bentley argued that willing another person's good might entail correcting perceived moral flaws in the other under an Aristotelian model of civic friendship. But under the pseudo-Aristotelian model of civic friendship, the citizen does not presume to know the good for the fellow-citizen and thereby attempt to bring the fellow-citizen into conformity with some particular conception of virtue. Instead, the citizen acts so as to provide the fellow-citizen with things useful for satisfying the other's rational desires. This arrangement privileges no particular understanding of good character or the comprehensive good and should therefore avoid the imposition of values that Bentley fears.

This arrangement in which civic friends attempt to secure for their fellow-citizens the primary goods necessary to carry out their life plans also addresses Kahane's concern. Kahane argued that Aristotelian civic friendship focuses on a characteristic of an individual and depends on cultivating a uniform conception of the good across society – regardless of what conception of the good and which characteristics were privileged. The pseudo-Aristotelian model of civic friendship does require a single uniform *concept* of the good, but without requiring a uniform *conception* of the good.³⁰ The thin theory of the good invoked by the pseudo-Aristotelian model still allows for wide diversity in comprehensive understandings of the good, with only a bare minimum of sameness required amongst the body of citizens. While there is still some requirement for sameness to support civic friendship on this model, this level of requirement should not constitute a reductio, at least for the Rawlsian liberal who appears to limit this sort of concern to fellow-citizens holding reasonable comprehensive doctrines. Meanwhile, unreasonable comprehensive doctrines are treated as a disease to be contained.³¹ This containment will enforce sufficient sameness for pseudo-Aristotelian civic friendship to flourish by keeping all comprehensive doctrines reasonable.

While this does entail a level of concern for the character of others, the concern is limited to public aspects of the fellow-citizens' character and not the others' comprehensive moral character. The concern extends only to whether or not the fellow-

³⁰ A concept, in Rawls' usage, is very rudimentary, as opposed to a conception which contains fuller detail. For example, Rawls employs the concept of a person as "someone who can partake in, or who can play a role in, social life, and hence exercise and respect its various rights and duties." (PL, 18) The conception of a person he employs is much more detailed. Rawls shows the stronger nature of a conception when he writes, "It is this political conception of persons, with its account of their moral powers and higher-order interests, [...] that provides the requisite background for specifying citizens' needs and requirements." John Rawls, *Political Liberalism* (Columbia University Press, 2005), 178.

³¹ Rawls, *Political Liberalism*. p64, n19.

citizen's comprehensive doctrine is a reasonable comprehensive doctrine.

Schwarzenbach states her requirements for a modified theory of civic friendship,

A central difference between ancient and modern civic friendship may be stated thus: the latter – considering the fact of modern pluralism – must legitimately operate via a doctrine of individual rights as well. That is, the legitimate care and concern democratic citizens reveal to one another today can no longer be dogmatically imposed, but operates within the range of recognized legitimate differences in religion, culture and moral sensibility; it refers to an overlapping minimal conception of citizens' good.³²

By restricting the locus of concern for one's fellow-citizens' character, this model makes the liberal deviations from Aristotelian civic friendship that Schwarzenbach maintained were necessary under conditions of pluralism.

The pseudo-Aristotelian model of civic friendship also reduces the chances of citizens failing to recognize well-wishing on the part of their fellow-citizens. The extensions of well-wishing present in society become much more uniform when the understanding of the good employed becomes more general. Actions calculated to increase the quantities of primary goods available to fellow-citizens may sometimes fail and result in no increase or even a decrease in primary goods. But such instances should be relatively rare in a society that takes reasonable measures to deliberate on policy decisions. Even when such errors are made, the background culture's broad employment of the thin theory of the good will help to reassure citizens that their fellow-citizens will the good for them. Citizens' fulfillment of 2** increases their chances of understanding that a misguided extension of well-wishing still corresponds to an intension of well-wishing. The background culture of well-wishing according to the thin theory of the good should provide a sufficient model for new generations to recognize the public

³² Sibyl A. Schwarzenbach, "Fraternity and a Global Difference Principle: A Feminist Critique of Rawls and Pogge." *International Politics* 48, no. 1 (2011): 28–45. p5.

political character of their peers and adopt it for themselves on a broad enough basis for society to remain stable. The benefits of the background institutions become apparent to members of the new generation as established members of society wish the youth well. The minimal commitment to a uniform concept of the good, though not to a uniform conception of the good, allows civic friendship thus modeled to flourish across a pluralistic society with few limitations, though one salient limitation must be addressed. Willing the satisfaction of other's rational desires is generally acceptable to people of a wide range of comprehensive doctrines, requiring participants to accept only the minimal commitment that it is not forbidden to will the satisfaction of the rational desires of others without first knowing the content of those desires. As such, models of citizenship based on this pseudo-Aristotelian civic friendship may be able to garner support of an overlapping consensus of reasonable comprehensive doctrines. Yet the commitment to wish fellow-citizens well while understanding the content of "well" according to the thin theory of the good minimizes, rather than eliminates, the mutual recognition issue. We ought not, however, assume that people of all comprehensive doctrines will be able to will the satisfaction of the rational desires of their fellow-citizens without any further specification of the content of those rational desires.

The Commitment Required

Civic friendship on the pseudo-Aristotelian model requires a minimum commitment from the citizens to a broad concept of the good. This model of civic friendship requires citizens to will for their anonymous fellow-citizens the satisfaction of their rational desires. As such, civic friendship thus understood requires that citizens find it permissible to will for their fellow-citizens that their rational desires be satisfied,

without specifying the content of those desires any further than that the desires are rational.

What aspects of a comprehensive doctrine might lead adherents to condemn willing the satisfaction of unspecified rational desires? One historical type of belief which could support this conclusion is the belief that a moral order exists prior to humanity, in the sense that right and wrong do not depend on individual human persons and their desires. According to such doctrines, this moral order, rather than the terms that individuals agree upon, determines what constitutes right action towards others. Such a belief does not entail condemnation of willing the satisfaction of unspecified rational desires, however. The comprehensive doctrine would need to further include the idea that humans are better off adhering to the dictates of the moral order than freely adopting rational desires contrary to the moral order and then satisfying (or attempting to satisfy) them. A comprehensive doctrine that has these two features, in addition to an injunction that adherents ought to care for others by willing that the others will be best off according to the adherent's comprehensive (rather than mere political) understanding, will have adherents that cannot participate in civic friendship according to the pseudo-Aristotelian model I have outlined.

For now, I say only that Rawlsian political liberalism requires the ascription of positive value to the satisfaction of unspecified rational preferences. Civic friendship, on this pseudo-Aristotelian account, requires less. It requires only that negative value not be ascribed to the satisfaction of unspecified rational desires. Citizens who assign no value at all to the satisfaction of others' unspecified rational desires may still engage in civic friendship under this model, motivated by the desire to achieve certain social goods.

Only those who are opposed to willing for their fellow-citizens the satisfaction of rational desires prior to a specification of the content of those desires will be excluded from participating in civic friendship on this model. Historically, however, some members of society do adhere to comprehensive doctrines that do not grant neutral or positive value to the satisfaction of rational desires of humans prior to specifying those desires. In the next chapter, I will look at a commitment to a broad concept of the good, similar to the one employed in pseudo-Aristotelian civic friendship, as it appears in Rawls' political liberalism.

CHAPTER THREE

The Rawlsian Value Assumption and Metaphysical Neutrality

And this end must have real value in their [Gaius' and Titius'] eyes. To abstain from calling it good and to use, instead, such predicates as 'necessary' or 'progressive' or 'efficient' would be a subterfuge. They could be forced by argument to answer the questions 'necessary for what?', 'progressing towards what?', 'effecting what?'; in the last resort they would have to admit that some state of affairs was in their opinion good for its own sake.

—CS Lewis, *The Abolition of Man*

In this chapter, I argue that Rawls' thin theory of the good is not metaphysically neutral, contrary to Rawls' claims that justice as fairness is political and not metaphysical. Rawls wants his theory of justice to be neutral with regards to substantive metaphysical claims that cannot be settled due to the burdens of judgment.¹ While Michael Sandel argues in *Liberalism and the Limits of Justice* that Rawls' argument to the original position depends on the Kantian conception of the person, I argue that Rawls' argument from the original position also makes objectionable metaphysical commitments. Rawls wants to make sure that the parties in the original position do not employ any doctrines that would be part of a comprehensive doctrine in order that the results of the thought experiment may possess objectivity. The parties in the original position require

¹ The burdens of judgment are a critical concept in understanding Rawls' work. After *Theory* Rawls introduces the term, and employs it afterward. In *Restatement*, Rawls writes "[...]the sources of reasonable disagreement – what I call the burdens of judgment – among reasonable persons are the many obstacles to the correct (and conscientious) exercise of our powers of reason and judgment in the ordinary course of political life." (Rawls, *Restatement*, 35) The burdens of judgment significantly impact our deliberation, "[...] many of our most important political judgments involving the basic political values are made subject to conditions such that it is highly unlikely that conscientious and fully reasonable persons, even after free and open discussion, can exercise their powers of reason so that all arrive at the same conclusion." (Rawls, *Restatement*, 36)

some theory of value by which to choose a set of principles of justice. To meet this requirement, Rawls gives the parties in the original position the thin theory of the good, which he argues is merely descriptive and therefore not objectionably metaphysical. I argue that the thin theory is insufficient to this task because Rawls must also attribute to these parties a belief that the satisfaction of their constituents' rational desires is worthy of choice. I then argue that this assumption is a metaphysical assumption in the sense that Rawls attempts to avoid. The argument that the assumptions made are *objectionably* metaphysical will follow in the next chapter.

Rawls admits that the parties in the original position need the thin theory of the good to make their decision, but claims that this does not entail that the parties have a theory of value which makes objectionable commitments and is thus consistent with the priority of the right over the good. I argue that the thin theory of the good must contain at least a minimum commitment to some value in order for the parties in the original position to make any decision. Without such a commitment, the argument *from* the original position (concluding that the parties chose the two principles of justice as fairness) cannot be sound. I begin by laying out the need for the parties in the original position to have some theory of the good to guide their decisions, specifically examining Rawls' case from *Political Liberalism* that the parties only need to recognize things as good in the very minimal sense required to give advice to another person on fulfilling their ends. I then argue, in parallel with Talbot Brewer's work in *The Retrieval of Ethics*, that some ascription of value, not merely usefulness, is required to give the parties in the original position motive to work towards the fulfillment of their constituents' just desires. If the thin theory of the good is purely descriptive as Rawls claims, then it is insufficient

to motivate the parties in the original position to act in such a way as to choose a set of principles. I enumerate the pieces of information from which Rawls claims we can deduce that the parties in the original position will choose the two principles of justice as fairness. I argue that none of these is sufficient to explain the choice made by the parties in the original position. Without a value assumption, the argument *from* the original position cannot be sound.

I then argue that the ascription of value by the parties in the original position constitutes a metaphysical assumption in the sense of metaphysical Rawls employs when he states that justice as fairness is political and not metaphysical. After explaining the objectionable sense of metaphysical in Rawls, I argue that these value ascriptions presupposed in Rawls' original position are of this sort, because they affirm answers to perennial philosophical questions. Thus Rawls' reasoning from the original position is not neutral between worldviews as he claims, but dependent on a contested set of metaphysical beliefs. This is to say that the reasoning Rawls would have us do after assuming the point of view represented by the original position is not neutral, but rather dependent on a controversial philosophical thesis. In the next chapter, I will examine this thesis further and argue that this commitment to a philosophically controversial position is not trivial.

The Need for the Thin Theory

Rawls contrasts justice as fairness with teleological theories of justice. His use of the term teleological is specific, referring to theories which presuppose some thick theory of the good for humans and proceed from that theory to an account of justice. For example, utilitarian theory is teleological in that it first defines the good (pleasure and the

absence of pain) and uses this to derive what is just. When utilitarianism condemns a social policy as unjust, it does so because the policy does not maximize the balance of pleasure over pain.² The conception of the good drives the pronouncement that the policy is unjust. Rawls' approach puts the right prior to the good in that he begins by searching for the principles of justice and then uses these principles to determine the complete theory of the good. But Rawls freely admits that the parties in the original position require some theory of the good to guide their decisions. Rawls addresses this need by equipping the parties with the thin theory of the good, also known as goodness as rationality, which recommends choices based on usefulness for fulfilling the desires of an agent. Rawls summarizes this theory, "to put it briefly, the good is the satisfaction of rational desire."³ The thin theory of the good is not a comprehensive theory of the good. It does not support all of our uses of the word good, only some of them.⁴ As such, starting with the thin theory of the good is consistent with the priority of the right according to Rawls.

The thin theory of the good has two purposes according to Rawls, "as what I referred to in [*A Theory of Justice*] as the thin theory of the good, goodness as rationality provides part of a framework serving two main roles: first, it helps us to identify a workable list of primary goods; and second, relying on an index of these goods, it enables us both to specify the aims (or motivation) of the parties in the original position and to

² Rawls has in mind here a specific type of utilitarianism. He writes, "I shall understand the principle of utility in its classical form as defining the good as the satisfaction of desire, or perhaps better, as the satisfaction of rational desire. This accords with the view in all essentials and provides, I believe a fair interpretation of it." John Rawls, *A Theory of Justice* (The Belknap Press of Harvard University Press, 1971), 25.

³ Rawls, *A Theory of Justice*, p93.

⁴ *Ibid.* p407.

explain why those aims (or motivations) are rational.”⁵ Rawls’ difference principle requires the identification of the least advantaged members of society, a task made possible by the account of primary goods.⁶ Generating the list of primary goods, especially goods such as self-respect, requires some theory of the good to guide the parties to include them. Rawls acknowledges this, “thus the initial definition of expectations solely by reference to such things as liberty and wealth is provisional; it is necessary to include other kinds of primary goods and these raise deeper questions. Obviously an account of the good is required for this; and it must be the thin theory.”⁷

Behind the veil of ignorance, the parties have very little to go on to guide their actions on behalf of their constituents. Rawls anticipates the challenge that the parties have too little information to go on and cannot make any progress given their lack of information about their constituents. Rawls defends the parties as making the best of a difficult situation, yet he is also cautiously optimistic that some progress can be made.

Nevertheless, the thin theory of the good which the parties are assumed to accept shows that they should try to secure their liberty and self-respect, and that, in order to advance their aims, whatever these are, they normally require more rather than less of the other primary goods. In entering into the original agreement, then, the parties suppose that their conceptions of the good have a certain structure, and this is sufficient to enable them to choose principles on a rational basis.⁸

While Rawls maintains that the thin theory of the good is sufficient for establishing the primary goods and choosing principles of justice, the theory is also modest and not sufficient for understanding all of our uses of the word good. Rawls tells

⁵ Rawls, *Political Liberalism*, p178.

⁶ Rawls, *A Theory of Justice*, p396.

⁷ *Ibid.* p396-397.

⁸ *Ibid.* p397.

us about the limitations of the thin theory, “but looking ahead to other questions yet to be discussed, a more comprehensive account of the good is essential. Thus the definition of beneficent and supererogatory acts depends upon such a theory. So likewise does the definition of the moral worth of persons.”⁹ The limitation is important. Regarding the conception of persons that is coupled with the thin theory of the good Rawls writes, “it is crucial here that the conception of citizens as persons be seen as a political conception and not as one belonging to a comprehensive doctrine.”¹⁰ Even if the political conception of citizens as persons is not directly used behind the veil of ignorance, the list of primary goods that it helps to generate will be used there and it is important that it not depend on any particular comprehensive doctrine. If this is the case, then the thin theory of the good must likewise avoid dependence on a comprehensive doctrine, since the thin theory of the good is explicitly employed behind the veil of ignorance. Thus we see that Rawls wants very much for the thin theory of the good not to depend on any comprehensive doctrine. If the parties in the thought experiment of the original position rely upon a comprehensive doctrine in their thinking, then Rawls is asking each of us, when we contemplate political matters from the point of view he argues we should take up, to also adopt that comprehensive doctrine. In turn, this would make justice as fairness metaphysical, not political.

Choiceworthiness

Rawls tells us that the parties in the original position must be guided by the thin (descriptive) theory of the good, along with the list of the primary goods (also dependent

⁹ *Ibid.* p397.

¹⁰ Rawls, *Political Liberalism*, p178.

on the thin theory of the good) so that we can explain their thinking. I argue that the thin theory of the good, in conjunction with the list of primary goods, is not sufficient to explain the rationality of the parties in the original position. The parties in the original position must choose a conception of justice from a slate of options. To explain their deliberation, some theory of justice must appear choiceworthy to the parties by appealing to some ascription of value.

Rawls does not think that he needs a robust sense of good for the parties in the original position to settle on his two principles of justice. He looks only for a descriptive, as opposed to a prescriptive, meaning of good to move his argument forward. That is to say, good is simply a way of describing a thing, akin to describing something as blue, without a weighty pronouncement that the thing holds a metaphysically esteemed position, or a position that is somehow particularly conducive to human flourishing. Rawls holds that goodness as rationality is a descriptive theory. He defines a descriptive theory as maintaining two theses. “First, despite the variation in criteria from object to object, the term ‘good’ has a constant sense (or meaning) that, for philosophical purposes, is of the same kind as that of other predicates normally counted as descriptive. [...] The other thesis is that the propriety of using the term ‘good’ in giving advice and counsel, and in expressions of commendation, is explained by this constant sense together with a general theory of meaning.”¹¹ This constant sense of the word that Rawls intends is the sense we employ when we give advice, saying things like “this is a good car,” or “that watermelon looks good.” Rawls maintains that these cases, unless situated in bizarre circumstances, do not make major metaphysical claims. Rawls completes his point, “the

¹¹ Rawls, *A Theory of Justice*, 405.

constant descriptive sense together with the general reasons why persons seek out the views of others explain these characteristic uses of ‘good.’ At no point must we appeal to a special kind of prescriptive or emotive meaning.”¹² All that is needed to explain these central uses of the word good is the thin theory of the good – that the thing has the features it is rational to want in a thing of its kind. There may be some uses of the word good not accounted for by this theory, but as noted previously, that is compatible with Rawls’ demands for the thin theory of the good.

It doesn’t seem that a strictly descriptive sense of good is sufficient to motivate the parties in the original position. In *The Retrieval of Ethics*, Talbot Brewer argues that the common belief/desire psychology used to explain the actions of agents is inadequate. In doing so, he argues that a desire for an outcome that renders some proposition true plus the belief that an action will bring about such an outcome is insufficient to explain why the agent subsequently undertakes to perform the action if a desire amounts to nothing more than a propensity to bring about such outcomes. Somewhere in an adequate explanation the object of choice must appear to be worthy of choice.

To be an agent is to set oneself in motion (or to try to do so, or to adopt the intention of doing so) on the strength of one’s sense that something counts in favor of doing so. That performing some action would bring about some state of affairs cannot intelligibly be regarded as counting in favor of performing the action unless one sees the state of affairs, or the effort to produce it, as itself good or valuable. Hence if we are to view persistent attempts to bring about some state of affairs as the doing of an agent, we must suppose that the agent sees something good or valuable about attaining or aiming at that state of affairs.¹³

Brewer’s point is that in order for an agent (such as the parties in the original position) to set herself into motion (such as agreeing to one conception of justice over another)

¹² *Ibid.* p406.

¹³ Talbot Brewer, *The Retrieval of Ethics*. Oxford University Press, 2009. p28.

requires an ascription of value to something. Regardless of whether the belief/desire model of psychology stands or falls, we cannot explain an individual's choice without claiming that the individual saw that option as worthy of choice.¹⁴

But this ascription of value cannot be grounded in the thin theory of the good. The thin theory of the good only recognizes something as good when that something serves to help the agent accomplish some pre-existing end. As such, neither the thin theory, nor anything else derived solely from the thin theory, can provide an ultimate end which the parties in the original position pursue when they choose which theory of justice to prefer. At best, the thin theory can point out means that will help achieve ends which are already deemed by the agent to be choiceworthy. Throughout *A Theory of Justice*, Rawls argues that the parties in the original position will determine that justice as fairness is more conducive to the fulfillment of their constituents' chosen life plans than principles based on utilitarianism. Thus, choosing justice as fairness would be described in the thin sense as being the better choice to accomplish the end of satisfying the constituents' chosen life plans. Yet Rawls' arguments that these principles will best fulfill their constituents' life plans does not explain why the representatives should see selecting justice as fairness as choiceworthy, because the thin theory of the good is merely descriptive. The assumption is limited – Rawls writes, “there is no inconsistency, then, in supposing that once the veil of ignorance is removed, the parties find that they have ties of sentiment and affection, and want to advance the interests of others and to

¹⁴ I may be understating the importance of this point – even if Rawls has a response to Brewer's argument, the parties in the original position are forced to take a side on an issue contended by philosophers – they must then side with Rawls in a debate about the nature of human choosing against Brewer. The parties in the original position then must be assumed to be convinced by Rawls' argument that Brewer is wrong. Such a conviction regarding the nature of human choosing may itself be a metaphysical commitment, and a commitment that Talbot Brewer may very well reject.

see their ends attained. But the postulate of mutual disinterest in the original position is made to insure that the principles of justice do not depend upon strong assumptions.”¹⁵ Yet the representatives’ choice can only be understood if they already see the fulfillment of their constituents’ life plans as a choiceworthy end. There must be another reason for the parties to adopt this initial set of ends, other than the thin theory of the good, that the parties in the original position pursue.

The assumption that fulfilling the rational desires of persons is choiceworthy can also be seen in another way. The parties in the original position arrive at the two principles, which, given their serial ordering, assigns zero weight to certain desires. Rawls writes, “[the priority of justice over efficiency and the priority of liberty over social and economic advantages] mean that desires for things that are inherently unjust, or that cannot be satisfied except by the violation of just arrangements, have no weight. There is no value in fulfilling these wants and the social system should discourage them.”¹⁶ This strongly implies that the social system acts to fulfill wants which are not unjust, treating such wants as having positive value. The parties in the original position would have no motive to choose principles which assign value to such desires unless the parties themselves ascribe value to the fulfilling of those desires. The fact that the parties find such a system worthy of choice demonstrates that the parties assume that fulfilling the rational desires of persons is of positive value. Since, as a purely descriptive theory, the thin theory of the good cannot ground such an ascription of value, we must

¹⁵ Rawls, *A Theory of Justice*, p129.

¹⁶ *Ibid.* p261.

investigate what other resources are available to the parties in the original position to explain their belief that fulfilling the desires of their constituents is worthy of choice.

CS Lewis argues that ascriptions of value must motivate action in *The Abolition of Man*. Lewis writes:

From propositions about fact alone no *practical* conclusion can ever be drawn. *This will preserve society* cannot lead to *do this* except by the mediation of *society ought to be preserved*. *This will cost you your life* cannot lead directly to *do not do this*: it can lead to it only through a felt desire or an acknowledged duty of self-preservation. The Innovator is trying to get a conclusion in the imperative mood out of premises in the indicative mood: and though he continues trying to all eternity he cannot succeed, for the thing is impossible.¹⁷

Rawls has taken pains to defend the thesis that the thin theory of the good is purely descriptive. This thin theory cannot supply the end for which the parties in the original position will move themselves to act when Rawls tells us they will choose the two principles of justice as fairness over other possible theories of justice. But Rawls' overall argument requires that individuals in the original position would choose his two principles of justice. The thin theory may serve for creating the list of primary goods, but it is insufficient to explain why the parties in the original position would choose *anything*, let alone the two principles of justice as fairness. The argument *from* the original position requires some value assumption.

The Deliberation of the Parties

For Rawls' argument, the parties in the original position must have some motive to choose the two principles of justice as fairness. But the limitations of the original position restrict the possible sources of the parties' ascription of choiceworthiness to the two principles. Rawls summarizes the sources to which the parties might appeal, "the

¹⁷ C.S. Lewis, *The Abolition of Man*. HarperCollins, 2001 p31-32.

veil of ignorance achieves this result [unanimous agreement] by limiting the parties to the same body of general facts (the presently accepted facts of social theory) and to the same information about the general circumstances of society: that it exists under the circumstances of justice, both objective and subjective, and that reasonably favorable conditions making a constitutional democracy possible obtain.”¹⁸ In order to understand why the parties in the original position would ascribe choiceworthiness to the two principles of justice as fairness, I proceed by enumerating the information available to the parties, as well as what Rawls tells us about their situation.

Rawls gives us a basic list of the premises supporting the parties’ conclusion. He writes, “we characterize the original position by various stipulations – each with its own reasoned backing – so that the agreement that would be reached can be worked out deductively by reasoning from [1] how the parties are situated and described, [2] the alternatives open to them, and [3] from what the parties count as reasons and [4] the information available to them.”¹⁹ This list does not exhaust the reasons available to the parties however. Rawls also writes,

To solve this problem [of not being able to decide] is one reason we introduced the idea of primary goods and enumerated a list of items falling under this heading. As we saw, these goods are identified by asking which things are generally necessary as social conditions and all-purpose means to enable citizens, regarded as free and equal, adequately to develop and fully exercise their two moral powers, and to pursue their determinate conceptions of the good.²⁰

The list of primary goods depends on the political conception of the person (as having the two moral powers and a determinate conception of the good) and the thin theory of the

¹⁸ Rawls, *Justice as Fairness: A Restatement*, p86-87.

¹⁹ *Ibid.* 17.

²⁰ *Ibid.* p88.

good.²¹ While I have already argued that the thin theory of the good cannot explain the parties' value ascription, both of these appear to belong on a list of the reasons available to the parties.

The parties in the original position are in a strictly regulated environment – there are only a few possible sources for their ascription of value to fulfilling the unspecified purposes of citizens. The list compiled above restricts these sources to the following:

1. The political conception of a person.
2. How the parties are situated and described.
3. The representation relationship.
4. The alternatives open to the parties.
5. What the parties count as reasons.
6. The information available to them.
7. The thin theory of the good.

The Political Conception of the Person Cannot Guide Choice

Since I have already argued that the thin theory of the good cannot ground the parties' ascription of choiceworthiness to fulfilling the rational desires of their constituents, I have moved it to the end of the list and will not reexamine it here. Instead, I proceed to examine the political conception of the person and whether or not this might assist in grounding the parties' value ascription. Rawls writes, "to conclude: given the conception of the person in justice as fairness, we say that the parties assume that, as

²¹ These two moral powers are "the capacity to understand, to apply, and to act from (and not merely in accordance with) the principles of political justice that specify the fair terms of social cooperation" and "the capacity to have, to revise, and rationally pursue a conception of the good." (Rawls, *Restatement*, 18-19)

persons with the two moral powers and a determinate complete conception of the good, citizens have, among other interests, certain religious, philosophical, and moral interests, and that the fulfillment of these interests must, if possible, be guaranteed."²² Here Rawls seems to think that the parties ascribe value to fulfilling their constituents' desires on the basis of the political conception of the person.

The two moral powers to which Rawls alludes are the capacity for a sense of justice and the capacity for a conception of the good.²³ This is part of understanding citizens as being free and equal persons. The broader understanding of citizens as free and equal persons also includes two ways in which citizens regard themselves as free. They are free in understanding that they have the capacity for a conception of the good²⁴ and also in that "they regard themselves as self-authenticating sources of valid claims".²⁵ Whatever determinate conception of the good a citizen possesses, she will regard her claims on societal institutions as being valid claims in light of her understanding of herself as a free and equal person.

Perhaps the representatives will choose a framework because that framework does the most thorough job of fulfilling valid claims on society by individuals within society. If the representatives are to view a citizen's claims as claims which must be fulfilled, then the representatives would need to know that those claims *are* valid claims. The political conception of the person does not grant that the citizen's claims *are* valid, only that the citizen views their claims as being valid. Knowing that the citizen views their claims as

²² *Ibid.* p104.

²³ *Ibid.* p18-19.

²⁴ *Ibid.* p21.

²⁵ *Ibid.* p23.

being valid however, is insufficient for the representatives to know that the citizen's claims are valid claims. In fact unreasonable citizens also regard themselves as self-authenticating sources of valid claims – yet the two principles of justice will not make any effort to fulfill those claims, because the claims, while viewed by the person making them as valid, are not in fact valid. This counter example shows that the citizen's view of her claims as being valid is insufficient for the representatives to know that the citizen's claims are valid. The source for the (potential) validity of these claims, be it correspondence with reality or based on an objective procedure, does not impact this result – the fact that the claimant views the claims as valid fails to guarantee that the claims are in fact valid. Since the representatives cannot know that the citizen's claims are valid, the representative cannot view a citizen's claims as claims which must be fulfilled. As such, the representatives cannot ascribe value to fulfilling the desires of their constituents on the basis of the freedom and equality of their constituents, nor on the basis of having the two moral powers and a determinate conception of the good.

The higher-order interests cannot guide choice. Another possibility is that the parties may find grounds for their choice in an entailment of the political conception of the person. Rawls holds that citizens have at least three higher-order interests which may suffice to motivate the parties in the original position. He writes:

We take moral persons to be characterized by two moral powers and by two corresponding highest-order interests in realizing and exercising these powers. The first power is the capacity for an effective sense of justice, that is, the capacity to understand, to apply and to act from (and not merely in accordance with) the principles of justice. The second moral power is the capacity to form, to revise, and rationally pursue a conception of the good. Corresponding to the moral powers, moral persons are said to be moved by two highest-order interests to realize and exercise these powers. [...] Since the parties represent moral persons, they are likewise

moved by these interests to secure the development and exercise of the moral powers.

In addition, I assume that the parties represent developed moral persons, that is, persons who have, at any given time, a determinate scheme of final ends, a particular conception of the good. Thus the model-conception defines moral persons as also determinate persons, although from the standpoint of the original position, the parties do not know the content of their conception of the good: its final ends. This conception yields a third interest that moves the parties: a higher-order interest in protecting and advancing their conception of the good as best they can, whatever it may be.²⁶

So there are three higher-order interests. 1) to realize and exercise the capacity to understand, apply and act from principles of justice, 2) to realize and exercise the capacity to shape and pursue their understanding of the good, and 3) to protect and advance that conception of the good.²⁷

The higher-order interests can be understood either as desires held by the constituents in the sense that they are overwhelmingly strong desires – desires so strong that they must, by their nature, be regulative of the constituent’s choices, or they could be understood as objective interests stating what is objectively best for the agent. In the former interpretation of the higher-order interests, the higher-order interests are just another interest of the constituent for purposes of motivating the actions of the representatives in the original position – the difference is a difference of magnitude rather than kind. This is insufficient – absent some particular interpretation of what it means to represent someone (see Representation Cannot Guide Choice below) this cannot guide

²⁶ John Rawls, “Kantian Constructivism in Moral Theory.” in *Collected Papers*, ed. Samuel Freeman (303–58. Cambridge, MA: Harvard University Press, 1999.) p312-313.

²⁷ The cited text is actually ambiguous with respect to what the three higher-order interests are. Instead, the first two may be 1) to realize the capacities to understand, apply and act from principles of justice and to realize the capacity to shape and pursue their understanding of the good and 2) to exercise these same two capacities. While I believe the interpretation given in the main text is the correct one, I do not need to defend this interpretation for my argument to proceed.

the parties' choice. Therefore if the higher order interests are to motivate the parties apart from the representation relationship, they must be objective facts about what is best for the constituents. This second approach to the higher-order interests runs contrary to Rawls' entire project in *Political Liberalism* and beyond because claims regarding the higher-order interests become truth claims about objective reality, rather than political assumptions drawn from our common stock of shared ideas. Therefore, we ought to reject this reading of the higher-order interests.

For the parties in the original position to be motivated directly by the higher-order interests on the second understanding, the parties must be persuaded to agree to a metaphysical claim about the nature of persons. Such a claim seems unlikely given that Rawls derives the higher-order interests, not from a claim that citizens *are* moral persons, but rather from the weaker claim that citizens *believe themselves to be* moral persons. This may be seen as Rawls begins his argument for the higher-order interests stating, “now the citizens of such a society *regard themselves as* moral persons and as having a conception of the good ...”²⁸ Therefore, we ought to understand the higher order interests in the former sense, as desires possessed by all actual persons that are strong enough to regulate those persons' behavior. The universality of these higher-order interests is important to Rawls' project, not because the parties can be expected to agree that these interests pursue objective value, but rather because all constituents can be safely assumed to have these interests.²⁹ If all actual persons have them, then each party

²⁸ *Ibid.* p312. Emphasis mine.

²⁹ Sandel appears to share this reading of Rawls' overall project. See his interpretation of Rawls' response to criticisms of the employment of the Kantian conception of a person, especially *Liberalism and the Limits of Justice* p189ff.

in the original position could assume that the constituent she represents has these interests and that these interests will regulate the choices of the constituent.

These interests, as Rawls describes them, are interests of moral persons – these are interests that constituents in the real world have. Rawls argues that the parties in the original position are moved by these interests because they *represent* moral persons who must have these higher-order interests. The parties in the original position are moved, not by the higher-order interests themselves, but by their representation relationship that they bear to one who has those higher-order interests. In the next sections we shall see that this representation relationship cannot ground an ascription of choiceworthiness by the parties in the original position.

The Situation and Description of the Parties Cannot Guide Choice

Rawls explains the situation and description of the parties in the original position, “in describing the parties we are not describing persons as we find them. Rather, the parties are described according to how we want to model rational representatives of free and equal citizens. In addition, we impose on the parties certain reasonable conditions as seen in the symmetry of their situation with respect to one another and the limits of their knowledge (the veil of ignorance).”³⁰ While the veil of ignorance is an important part of their situation, this device does not impart any reasons for the party to value certain states of affairs – indeed it strips away almost all possible reasons they might have. The veil of ignorance, as part of the situation the parties find themselves in, is the barrier which they need to overcome by finding some way to ground an ascription of value – not a source for value ascriptions.

³⁰ Rawls, *Justice as Fairness: A Restatement*, p 81.

One might think that the ascription of value made by the parties in the original position can be derived from the description given of the parties. Rawls suggests such an approach when he writes, “we include the necessary psychology in the description of the parties as rational representatives who are moved to secure the good of those they represent, as this good is specified by the account of primary goods.”³¹ Rawls elaborates on this point further in view of the problem of the parties reaching a definite conclusion. He writes, “as we saw, these goods are identified by asking which things are generally necessary as social conditions and all-purpose means to enable citizens, regarded as free and equal, adequately to develop and fully exercise their two moral powers, and to pursue their determinate conceptions of the good.”³²

Yet we are left here with a further question – how is it that the parties choose to represent their constituents by promoting the fulfillment of the things that the constituents in reality perceive to be good? That is, why help their constituents to get the things that the constituents want, rather than, for instance, getting for the constituents what is objectively good for the constituents? For example, perhaps what is good for the constituents is not to simply develop and exercise their moral powers, but rather to exercise their moral powers in the correct way while exercising these powers in the wrong way would be harmful. If this option cannot be ruled out by the parties in the original position, then the parties cannot agree that simply allowing their constituents to exercise their moral powers in any means whatsoever would be worthy of choice. Restricting the exercise of one’s powers to pursuing only those desires the pursuit and

³¹ *Ibid.* p83.

³² *Ibid.* p88.

fulfillment of which is compatible with like liberty for all will not make this worthy of choice. Even if pursuit of desires that fail to meet this Rawlsian criterion can be assigned lesser value, this falls short of bestowing positive value on pursuing those desires that meet the criterion as the value of pursuing those desires which fail may be negative. Ruling this option out would require the employment of controversial metaphysical theses by the parties in the original position and hence the parties cannot rule out this option. The thin theory of the good can explain why the parties might suggest that their constituents should desire the primary goods to pursue their ends or use the thin label “good”. But the thin theory, the only theory of the good available to the parties, cannot ground the parties’ ascription of choiceworthiness to *any* particular political framework let alone the one Rawls claims they will choose.

If the situation of the representatives behind the veil of ignorance deliberating about fair terms of cooperation cannot explain the value ascription made by the parties in the original position, then we will have to move further down the list of resources Rawls gives to them. In the situation as described, there are two salient options for how the decision of which good to pursue might be made: 1) Seeking the fulfillment of the desires of the constituent (rather than some other good) may be entailed by the relationship between the representative and the constituent or 2) it may be entailed by the imperative to deliberate and agree upon fair terms of cooperation that the parties receive.

Representation Cannot Guide Choice

Another reason the representative as described might ascribe value to the constituent’s rational desires is the representation relationship. I take this relationship to be something other than a relationship of identity. The things in the real world that have

desires and conceptions of the good are individual persons. According to Rawls, the parties in the original position are not individual persons.³³ Hence, no party in the original position could be identical to any individual constituent in the real world. Later Rawls minimizes the importance of the distinction between seeing the parties as actual members of society behind the veil of ignorance or as representatives, “in the original position we may describe the parties either as the representatives (or trustees) of persons with certain interests or as themselves moved by these interests. It makes no difference either way, although the latter is simpler and I shall usually speak in this vein.”³⁴ Rawls has already told us that all persons view themselves as self-authenticating sources of valid claims.³⁵ Yet if the representative takes on the perspective of a constituent, the representative will (from that perspective) view the claims made by the constituent as self-authenticating and valid. The representative is thereby released from the necessity of making an ascription of value – that job has already been done by the constituent. The rationality of the constituent in making the ascription of choiceworthiness is not important, that ascription may be undertaken with the full comprehensive doctrine of the constituent who is not subject to the veil of ignorance. The representative in the original position does not need to know why the constituent ascribed choiceworthiness to some end or what the end is, only that the constituent ascribed value to it. If the representative takes on the viewpoint of the constituent, the representative accepts the constituent’s

³³ Rawls, *A Theory of Justice*. p146.

³⁴ Rawls, “Kantian Constructivism in Moral Theory” p312.

³⁵ Rawls, *Justice as Fairness: A Restatement*, p23.

ascription of choiceworthiness as her own and has sufficient motivation to choose whatever principles of justice will best secure that end for the constituent.

Rawls gives several indications supporting this reading. He writes, “rather, it is because it is rational for them [the parties] as trustees, and so as responsible for citizens’ (unknown) determinate and complete good, to deliberate in this [cautious] way, given the overriding importance of establishing a public conception of justice that guarantees the basic rights and liberties.”³⁶ The fact that the parties are trustees makes caution rational for the parties – it would not be if not for the trustee relationship with their constituents. Rawls also tells us about the parties that, “their reasoning aims at selecting the principles of justice that best secure those persons’ [their constituents’] good, their fundamental interests, ignoring any inclinations that might arise from envy...”³⁷ And again, Rawls tells us that, “the parties, as representatives of free and equal citizens, act as trustees or guardians. Thus, in agreeing to principles of justice, they must secure the fundamental interests of those they represent.”³⁸ Perhaps most clearly he writes, “whenever circumstances are relevant to [the highest-order interests’] fulfillment, these interests govern deliberation and conduct. Since the parties represent moral persons, they are likewise moved by these interests to secure the development and exercise of the moral powers.”³⁹

An account on which the representative assumes the point of view of a constituent in making choices provides the motivation necessary for the agent in the original position

³⁶ *Ibid.* p107.

³⁷ *Ibid.* p84.

³⁸ *Ibid.* p84.

³⁹ Rawls, “Kantian Constructivism in Moral Theory” p312.

to choose justice as fairness over utilitarian principles of justice. But this requires that that the representative in the original position chooses to execute her duties of representation by adopting the viewpoint of the constituent. To fully understand the deliberation of the representative in the original position, we must further seek to understand why the representative chooses to execute her duties in this fashion, rather than some other fashion. To adopt that standpoint is to accept as veridical the individual's perception that the individual is a self-authenticating source of valid claims. In seeking to understand the reasoning of the parties in the original position, we now need a way to understand why the representatives choose to represent their constituents by adopting the viewpoint of their constituents. Again, the parties in the original position have chosen one course of action (representation by adopting the viewpoint of the constituent) above all others and so must have made an ascription of choiceworthiness.

If the Rawlsian chooses to maintain that the representatives choose adopting the viewpoint of their constituents, then there must be a reason the parties find this approach worthy of choice. To justify that ascription of choiceworthiness, the parties could appeal to a further property of that means of representation either as intrinsically worthy of choice or to such a property as merely good in the thin sense. But if they appeal to such a property as good only in the thin sense, they will again have to have some reason to choose that property to endorse and so on forever. Without means for generating ascriptions of choiceworthiness (a controversial metaphysical thesis), the parties cannot *choose* to fulfill their duties in any particular way. To avoid this regress without objectionable metaphysical commitments, the Rawlsian should reject the possibility that the parties in the original position choose to fulfill their duties as representatives in a

certain way and instead claim that representation of another party simply entails taking on that party's point of view.

If Rawls is to explain the reasoning of the parties in the original position without reference to a prescriptive theory of the good, he must hold that the parties do not face any choice about how to represent their constituents and that the only way for them to represent their constituents is by adopting the point of view of their constituents. Somehow, buried in the question posed to the representatives or in the concept of representation, taking the constituent's point of view is entailed. This assumption, however, also seems to have significant metaphysical implications because the parties in the original position are repeatedly asked to defend the "fundamental interests" of their constituents. But the fundamental interests of the constituents are not clear. The parties in the original position have no reason to assume that the ends chosen by their constituents align with their fundamental interests. What is more, under conditions of pluralism regarding significant moral issues, at least some of the citizens must misunderstand their actual, as opposed to perceived, fundamental interests. The parties, unable to employ ideas that are not common to a diverse population of real-world citizens cannot appeal to an understanding of what sort of things their constituents are and cannot recognize their constituents' fundamental interests.

Rawls differentiates justice as fairness from other doctrines,

Consider again the idea of social cooperation. Let's ask: how are the fair terms of cooperation to be determined? Are they simply laid down by some outside agency distinct from the persons cooperating? Are they, for example, laid down by God's law? Or are these terms to be recognized by these persons as fair by reference to their knowledge of a prior and independent moral order? For example, are they regarded as required by natural law, or by a realm of values known by rational intuition? Or are these terms to be established by an undertaking among these persons themselves in the light of what they regard as

their mutual advantage? Depending on which answer we give, we get a different conception of cooperation. Since justice as fairness recasts the doctrine of the social contract, it adopts a form of the last answer: the fair terms of social cooperation are conceived as agreed to by those engaged in it, that is, by free and equal persons as citizens who are born into the society in which they lead their lives.⁴⁰

The decision to pursue the rational advantage of the parties' constituents is not uncontested. Recognition of the constituents' rational advantage does not render pursuing this advantage worthy of the parties' choice. There are other comprehensive doctrines that give alternative answers to the question of how the fair terms of cooperation are determined than the contractarian answer that Rawls puts forth. To assume that the only way for the representatives to properly represent their constituents is to adopt the viewpoint of the constituents and pursue their rational advantage is to have already given a contractarian answer to a question that remains contentious. This picture of representation is not a part of the fund of common ideas agreed upon by all members of society as clearly demonstrated by the background of ongoing dialogue against which Rawls makes his claims. The natural law tradition, rational intuitionism and belief in God as the source of justice in a Platonic sense all continue to find nontrivial numbers of adherents in our background culture. As such, for the parties in the original position to depend upon the contractarian understanding of their relationship to their constituents or the question being posed to them – rather than a competing understanding which would lead to very different deliberations – saddles justice as fairness with an assumption drawn from a comprehensive doctrine, rather than from a political conception or a shared fund of broadly accepted facts. If the original position contains such controversial

⁴⁰ John Rawls, "Justice as Fairness: Political Not Metaphysical." In *Collected Papers*, ed. Samuel Freeman, (388–415. Cambridge, MA: Harvard University Press, 1999) p399. See also Rawls, *Political Liberalism*, p22-23 and p97, and *Justice as Fairness: A Restatement* p14-15.

assumptions, it is not the right position from which to determine matters of basic justice and the argument *to* the original position is unsound. Therefore, the Rawlsian must reject the representation relationship as the basis for the parties' ascription of choiceworthiness to selecting justice as fairness from among competing conceptions of justice.

Choiceworthiness from an Imperative?

I can already hear the frustrated Rawlsian replying to this along the lines, “why do the parties need another motive to choose an answer? They have been given an imperative and there are no contentious metaphysical assumptions made if we assume that they find responding to that imperative, to the best of their ability, choiceworthy.” It is possible that the parties in the original position have a very simple doctrine of choiceworthiness that doing what they have been tasked with doing, to the best of their ability, is choiceworthy. The parties, without further guidance as to what is choiceworthy consulted their information and found that there was no consensus to be had when they searched for an answer to the question based on claims to objective truth. Given the burdens of judgment, the parties simply do the best they can to fulfill their mandate to find the best theory of justice among the list of potential answers presented to them for consideration with what they have.

The burdens of judgment, agreed upon as fact by real people in the real world and thus transmitted to the parties in the original position, serve to dissuade the representatives from pursuing answers based on disputed truth claims. Note that this does not rule out doctrines that make claims to absolute truths. For example, principles of justice based on a version of utilitarianism are entertained quite seriously by the parties in the original position, even though these principles are teleological, that is, based on a

prior understanding of the good for mankind.⁴¹ It is at least epistemically possible for the parties in the original position that the principles of utilitarianism will appear best. Yet the parties cannot choose the best theory of justice by appeal to objective truth. Without the ability to identify a theory of justice as choiceworthy based on objective truth, the parties in the original position might instead identify choiceworthiness by some other criterion. For example, Rawls portrays the parties as choosing based on expected outcomes. But adopting any other criterion in place of objective truth will commit the parties to the assumption that the property picked out by this criterion is in fact valuable.⁴²

For example, assume that the parties in the original position do accept as worthy of choice the idea that their constituents are best represented by adopting the constituents' points of view. A representative might choose this method of representation because the ends pursued were chosen in a certain way – they were chosen by the constituent. Of course, there might be other reasons for the representative's choice instead, those reasons are not important. What is important is the fact that the representative chose some criterion to identify fulfilling the imperative in this manner as the *best* that she could do. Rawls might suggest objectivity as the criterion. Regardless of the criterion chosen, if the imperative to choose principles is to motivate the parties to action, the representative must choose some controversial criterion to judge when she has fulfilled her mandate to

⁴¹ Indeed, utilitarian theories of justice seem to be the primary alternative considered by the parties when Rawls argues that the parties will choose justice as fairness in *A Theory of Justice*.

⁴² One may well notice that for the parties to choose a theory of justice on the basis that it is the theory of justice which describes the world as it really is also employs a contentious metaphysical assumption. A pragmatist situated in the original position could well find that being the one true theory of justice does not count as sufficient reason for it to be the theory of justice selected. Regardless, the possibility of choosing by appeal to truth is not open to the parties in the original position given the burdens of judgment.

the best of her ability. That is, if the imperative motivates the parties, then the controversial interpretation of the imperative means the original position is not the correct position for determining matters of basic justice and the argument *to* the original position is unsound.

The Alternatives Open to the Parties Cannot Guide Choice

Rawls describes for us the work done by the list of alternatives available to the parties in the original position. He writes, “to proceed: we assume that the parties reason by comparing alternatives two at a time. They begin with the two principles of justice and compare those principles with the other available alternatives on the list. If the two principles are supported by a stronger balance of reasons in each such comparison, the argument is complete and those principles are adopted.”⁴³ The list of alternatives presented to the parties in the original position makes the overall choice of principles of justice into a series of contrastive decisions. Rawls subsequently argues that the two principles of justice are the Condorcet winner, beating every other option on the list in a one to one comparison.

Yet while the list of alternatives helps to frame the way in which the parties approach the choice, it does not give the parties a source of value which might determine which outcome is worthy of being chosen by the parties. If it were the case that the list of alternatives provided a source of value, then enumerating the members of the list of alternative conceptions of justice would be a topic of high importance for Rawls. In fact, Rawls makes no attempt to enumerate the members of the list of alternatives.

⁴³ Rawls, *Justice as Fairness: A Restatement*, p95.

In *Theory* he writes, “I shall consider for the most part the choice between the two principles of justice and two forms of the principle of utility. Later on, the comparisons with perfectionism and mixed theories are discussed.”⁴⁴ But Rawls thinks we do not need to be more specific to reach his conclusion, “now admittedly this is an unsatisfactory way to proceed. It would be better if we could define necessary and sufficient conditions for a uniquely best conception of justice and then exhibit a conception that fulfilled these conditions. For the time being, however, I do not see how to avoid rough and ready methods.”⁴⁵ While Rawls does produce a specific list of the alternatives he will use in *Theory*, he does not seem to think the choice of those alternatives is privileged. He writes,

If we change the list, the argument will, in general, have to be different. A similar sort of remark applies to all features of the original position. There are indefinitely many variations of the initial situation and therefore no doubt indefinitely many theorems of moral geometry. Only a few of these are of any philosophical interest, since more variations are irrelevant from a moral point of view. We must try to steer clear of side issues while at the same time not losing sight of the special assumptions of the argument.⁴⁶

If argument over the exact contents of the list of alternatives is a side issue, then we can safely conclude that the content of the list of alternatives does no significant work in convincing the parties in the original position that one particular alternative is worthy of their choice. Furthermore, if Rawls felt that his conclusion depended on the list of alternatives, his conclusion would be very weak indeed – any suggestion of a theory which ought to be included would undermine his argument without actually having to

⁴⁴ Rawls, *A Theory of Justice*, p123.

⁴⁵ *Ibid.* p123.

⁴⁶ *Ibid.* p126.

argue the merits of such a theory to the parties in the original position. Instead, Rawls writes in the *Restatement*, “we do not claim that the two principles would be agreed to from a complete, or any possible, list. To claim that would be excessive and I attempt no general argument. The two comparisons we will discuss are, then, but a small part of the argument that would be required to provide a reasonably conclusive argument for the two principles of justice.”⁴⁷

What the Parties Count as Reasons Cannot Guide Choice

Regarding what the parties in the original position count as reasons, Rawls writes,

To illustrate regarding appropriate restrictions on reasons: if we are reasonable, it is one of our considered convictions that the fact that we occupy a 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. [...] To model this and other similar convictions, we do not let the parties know the social position of the persons they represent. The same idea is extended to other features of persons by the veil of ignorance.⁴⁸

Rawls refers to this restriction on reasons as public reason. One can see that public reason will not answer any questions regarding how the parties in the original position come to make any value ascriptions that they make. If what the parties count as reasons is to help explain their ascription of value, then what they count as reasons must itself provide a reason to make a decision. But public reason is a restriction on reasons – it serves to prevent certain types of potential reasons from being employed by the parties in the original position. Public reason does not introduce a new set of reasons – while it may help eliminate competing reasons the parties might entertain (thus moving Rawls’

⁴⁷ Rawls, *Justice as Fairness: A Restatement*, p95.

⁴⁸ *Ibid.* p18.

argument from the original position forward) it cannot explain why the parties *do* find something worthy of choice.

The Information Available to the Parties Cannot Guide Choice

Rawls enumerates the information made available to the parties in the original position. “The fact of reasonable pluralism is the first of five such facts that are especially important in justice as fairness.”⁴⁹ He continues, “a second and related general fact is that a continuing shared adherence to one comprehensive doctrine can be maintained only by the oppressive use of state power. [...] Let us call this the fact of oppression.”⁵⁰ “A third general fact is that an enduring and secure democratic regime, one not divided by bitter doctrinal disputes and hostile social classes, must be willingly and freely supported by at least a substantial majority of its politically active citizens.”⁵¹ “We add, then, a fourth general fact: that the political culture of a democratic society that has worked reasonably well over a considerable period of time normally contains, at least implicitly, certain fundamental ideas from which it is possible to work up a political conception of justice suitable for a constitutional regime.”⁵² Finally, Rawls tells us that, “a fifth and last general fact may be stated as follows: that many of our most important political judgments involving the basic political values are made subject to conditions such that it is highly unlikely that conscientious and fully reasonable persons, even after free and

⁴⁹ *Ibid.* p33.

⁵⁰ *Ibid.* p34.

⁵¹ *Ibid.* p34.

⁵² *Ibid.* p34-35.

open discussion, can exercise their powers of reason so that all arrive at the same conclusion."⁵³

These five facts are insufficient to explain why the parties in the original position would ascribe value to the fulfilling of the unspecified desires of rational agents. The facts of pluralism and oppression attempt to describe a historical generalization – something Rawls believes all have learned from millennia of history. The third fact, a fact about stability, appears to be a similar generalization of our experiences with what can be learned from the failures of the past. These three facts are facts about the relationships between events, not normative statements that could provide grounds for an ascription of choiceworthiness.

The fourth and fifth facts do have at least some reference to value. The fourth fact applies to only those democratic societies that have worked reasonably well – value judgments will be needed to understand which societies have worked well or poorly. But the fourth fact does not provide a means for adjudicating disputes about which democratic societies have worked well – that judgment appears to be left up to the reader's judgment, presumably subject to the restraints of public reason. The fifth fact, of the burdens of judgment, gives us caution against insisting on a particular conception of the good to dominate our political society. The burdens of judgment underpin Rawls' case for the fact of reasonable pluralism. Some understanding of the good may be necessary to determine which of our political judgments are our most important political judgments. Yet the facts Rawls provides to the parties do not supply the standards by which political judgments are to be understood as important ones. If anything, this fact

⁵³ *Ibid.* p36.

should caution the parties in the original position against making new ascriptions of value, because their constituents in the real world cannot be expected to arrive at similar conclusions about which things are and are not of positive value. The burdens of judgment cannot explain why the parties in the original position ascribe value to fulfilling the unspecified desires of their constituents.

While Rawls appears to close his list with exactly five facts, there may be other items not included on this list as well. Rawls writes,

In presenting the argument for the two principles we have to refer to the parties' general knowledge of social theory and human psychology. But how is this knowledge specified? It must be settled by you and me as we set up justice as fairness. It is up to us to say what the parties are to know in view of our aims in working out a political conception of justice that can be, we hope, the focus of a reasonable overlapping consensus and hence a basis of public justification.⁵⁴

This statement allows the parties in the original position to have any additional information we want them to have as we discuss fair terms of social cooperation, including a theory of value that will motivate the parties' ascription of choiceworthiness to the two principles of justice. Yet Rawls must be careful regarding the items that are entered in the original position in this way – the items thus imparted to the parties in the original position must not prevent reasonable persons from joining in an overlapping consensus focusing on the two principles of justice as fairness. As charitable interpreters, we ought not construe Rawls' argument to depend on unstated premises smuggled in, rather than openly declared, through a loophole such as this.

I have argued that the parties in the original position must make some ascription of value to choosing the two principles of justice vis-à-vis the presented alternatives in order for them to choose the two principles as Rawls argues they will. That ascription of

⁵⁴ *Ibid.* p89.

value cannot be motivated by the thin theory of the good which Rawls allows the parties because the thin theory only identifies things that would be useful for the constituents to carry out their chosen ends. Nor can it be motivated by any of the other sources Rawls presents for our consideration, such as the higher-order interests, the description and situation of the parties, the options presented to them, what the parties count as reasons or the information given them. Rawls has made the silent assumption that things identified as useful to the constituents (and thus good in the thin sense) are also good in the more robust sense that might motivate the parties in the original position to choose a particular set of principles in order to obtain this good. I now argue that this assumption is a metaphysical assumption, in Rawls' sense of metaphysical employed in the phrase "political, not metaphysical." In the next chapter, I argue that this value assumption must not be added through these means, because it would prevent reasonable people from joining in an overlapping consensus focusing on the two principles of justice as fairness. If the value assumption is strong enough to make the argument *from* the original position work, it is too strong for the argument *to* the original position to be sound.

Rawls' Sense of Metaphysical

For justice as fairness to work as Rawls hopes, real people must be able to enter into the original position. That is, they must be able to adopt the point of view of one of the parties behind the veil of ignorance and think through which set of principles of justice they should employ. Ultimately, when individuals assume this point of view they need to arrive at an overlapping consensus on a political conception of justice. They do not need to agree on the intensional content of justice as fairness - the intensions of political conceptions may vary greatly. Notably, they do not need to agree on how best to

achieve these results, but an overlapping consensus on the ends must result from people adopting the point of view of the original position.⁵⁵ Metaphysical agreement is not required, but the conceptions need to agree in extension, that is, agree on the effects of the institutions that a well-ordered society ought to feature. Diverse individuals may justify this political conception from many different comprehensive doctrines and no comprehensive agreement is required. Indeed, such metaphysical agreement cannot be reached by any acceptable means, according to Rawls. Therefore, my conclusion here is limited – I do not make any claims about political liberalism generally, or that the extension of justice as fairness cannot be the subject of an overlapping consensus, but rather, that justice as fairness requires the representatives in the original position have certain contested value commitments.

Rawls instructs us, “conceptions of justice are to be ranked by their acceptability to persons so circumstanced [in the correct position, i.e. the original position]. Understood in this way the question of justification is settled by working out a problem of deliberation: we have to ascertain which principles it would be rational to adopt given the contractual situation.”⁵⁶ I have argued that the lines of thinking undertaken by the parties in the original position (intended to determine the acceptability of various conceptions of justice to those parties) depend on an assumption regarding value. The representatives, because they choose to pursue the satisfaction of the rational choices of their constituents, must view the satisfaction of said desires as valuable absent any specification of the content of those desires. This is reminiscent of Rawls’ statements

⁵⁵ See Rawls, *Political Liberalism*, p 150-154.

⁵⁶ Rawls, *A Theory of Justice*, p 17.

that, on his view, the fair terms of cooperation are determined by the agreement of the parties, rather than being “simply laid down by some outside agency distinct from the persons cooperating.”⁵⁷ The answers Rawls thinks the representatives in the original position will reach depend on giving those representatives a certain set of values – they must find the satisfaction of their constituents’ rational preferences worthy of choice. They choose the two principles of justice *because* those two principles give the highest expectation value for their constituents to satisfy their rational desires. I now seek to argue that this value ascription makes justice as fairness, contra Rawls, metaphysical and not merely political.

My claim comes with complications. In defending his claim that political liberalism is not metaphysical, Rawls writes, “part of the difficulty is that there is no accepted understanding of what a metaphysical doctrine is.”⁵⁸ Rawls also acknowledges the quantity of potential claims that his thesis is metaphysical and writes more than once, “to rebut claims of this nature [claims that a metaphysical doctrine of the person is presupposed] requires discussing them in detail and showing that they have no foothold. I cannot do that here.”⁵⁹ This makes my task difficult – since Rawls would not, and cannot, defend himself against such a charge, I must endeavor to be as generous as possible. Rawls’ claim that justice as fairness is political and not metaphysical should not be taken as a trivial claim. He has given us some ideas of what he means when he uses the word metaphysical in this way.

⁵⁷ Rawls, “Justice as Fairness: Political not Metaphysical” p399.

⁵⁸ *Ibid.* p403, n22.

⁵⁹ *Ibid.* p 403. Also *Political Liberalism* p29.

Rawls explains his sense of what it is for a claim to be metaphysical:

One might say, as Paul Hoffman has suggested to me, that to develop a political conception of justice without presupposing, or explicitly using, a metaphysical doctrine, for example, some particular metaphysical conception of the person, is already to presuppose a metaphysical thesis: namely, that no particular metaphysical doctrine is required for this purpose. One might also say that our everyday conception of persons as the basic units of deliberation and responsibility presupposes, or in some way involves, certain metaphysical theses about the nature of persons as moral or political agents. Following the method of avoidance, I should not want to deny these claims. What should be said is the following. If we look at the presentation of justice as fairness and note how it is set up, and note the ideas and conceptions it uses, no particular metaphysical doctrine about the nature of persons, distinctive and opposed to other metaphysical doctrines, appears among its premises, or seems required by its argument. If metaphysical presuppositions are involved, perhaps they are so general that they would not distinguish between the distinctive metaphysical views – Cartesian, Leibnizian, or Kantian; realist, idealist, or materialist – with which philosophy has traditionally been concerned. In this case, they would not appear to be relevant for the structure and content of a political conception of justice one way or the other.⁶⁰

I read this passage as presenting two criteria for a political conception to be metaphysical. Firstly, the political conception must either cite as a premise, or its argument must require some particular metaphysical doctrine about the nature of persons that sets it against other metaphysical doctrines. But I have argued that the parties in the original position do in fact need a thesis about the nature of persons – specifically that pursuing the satisfaction of their rational desires is worthy of choice. This is a thesis about the nature of persons because it highlights at least one thing that is good for humans – the fulfillment of their unspecified rational desires. Contrast this with a thesis that assigns no value to the fulfillment of rational desires unless those desires are specified, leading to ambivalence about fulfilling such desires, or with a thesis that assigns a negative value, motivating the parties to choose so as to prevent the pursuit and/or fulfillment of these

⁶⁰ *Ibid.* p403-4, n22.

desires. Such theses may (after specification) assign a positive value to those desires that meet certain qualifications and a negative (or zero) value to those that fail to meet those qualifications, but assign no value until after the desires are specified. The second criterion for a political conception to be metaphysical is that it would distinguish between major metaphysical views.

This second criterion does not merely restate the first criterion. Rawls is not afraid of contradicting major systems of thought in his setup for the original position. After all, he presumes that the parties in the original position are aware of the burdens of judgment, recognizing that properly functioning humans placed in similar circumstances cannot be expected to arrive at the same answer in reflective equilibrium as a result of their investigations of the good for mankind. Some hold that the existence of contingent reality entails theism and that the existence of contingent reality could be established without any particular appeal to a specific comprehensive doctrine. Some Christians respond to the problem of divine hiddenness by denying the premise that non-culpable non-belief exists.⁶¹ That is to say that any instance of an individual not believing in God (in at least some form) must be due to some blameworthy defect on the part of the individual – a view contradicted by the burdens of judgment. Accepting the thesis that divine hiddenness is a problem for Christian theists (especially the premise that non-culpable non-belief exists) may have significant ramifications for a person's comprehensive doctrine. Yet in spite of this, the burdens of judgment which Rawls presumes the parties in the original position accept as factual, does not (per Rawls) make

⁶¹ A sample formulation of the problem of divine hiddenness might look like this: Some non-belief is non-culpable and all non-belief is punished by God, then some non-culpable non-belief is punished by God.

the burdens of judgment a metaphysical thesis. This shows the importance of the criterion about major metaphysical positions.

The second criterion states that theses which are metaphysical in Rawls' sense would "distinguish between the distinctive metaphysical views – Cartesian, Leibnizian, or Kantian; realist, idealist, or materialist."⁶² The assumption that the fulfillment of human beings' unspecified rational desires is a good distinguishes between metaphysical views examined and debated by some of the most historically influential schools of philosophy. Rawls adopts a principle stating that "other things equal, human beings enjoy the exercise of their realized capacities, and this enjoyment increases the more the capacity is realized, or the greater its complexity".⁶³ He calls this principle the Aristotelian Principle and in doing so attests to the antiquity of philosophical debate over the nature of the good for humans.

Historically, Aristotle's ethics are significant because there is a certain way that people are *supposed* to be, an objective good for individual people to conform to. Epicureans held that the good was pleasure, and that rather than a way people were supposed to be so as to realize their nature, there was only a pragmatically best way to be in order to experience the greatest balance of pleasure over pain. Atomists argued that there were no gods and no afterlife, and for this reason, people ought to pursue the satisfaction of their desires. Stipulatively assigning value to the satisfaction of rational human desires apart from the content of those desires would settle significant disputes between these philosophers – it requires the existence of good in the simple satisfaction

⁶² *Ibid.* p404, n22.

⁶³ Rawls, *A Theory of Justice*, 426.

of desires, apart from any conformity or non-conformity of those desires to an objective good. Not only might it settle certain questions of how man ought to act, but because the prescribed actions are entailed by the underlying metaphysics, it would select among the metaphysical views as well. While Rawls' list of metaphysical doctrines does not include references to such ancient philosophers, to exclude the ancients' metaphysical doctrines from the set of views "with which philosophy has traditionally been concerned" would impoverish philosophy beyond reason. Since the commitment of the parties in the original position to ascribe value, whether directly or indirectly, to the satisfaction of the rational preferences of humans would decide such traditional philosophical questions, the assumption also satisfies the second criterion Rawls gives for an assumption to be metaphysical.⁶⁴

As indicated above, Rawls is aware that some metaphysical commitments might still be present in justice as fairness. He recognizes that perfection in this dimension might never be obtained, stating, "to secure this [informed and willing political] agreement we try, *so far as we can*, to avoid disputed philosophical, as well as disputed moral and religious, questions."⁶⁵ To be certain, some metaphysical assumptions are acceptable, even expected in any theory of justice. Few would count it against Rawls' theory of justice that it presupposes the existence of other minds, thereby dismissing solipsism – even though rejecting solipsism answers significant questions that distinguish between world views.

⁶⁴ The essentialist idea that there is a set of right purposes for humans that are all-things-considered binding on them, regardless of their desires and chosen goals still has some currency in philosophy. Examples include Mark Murphy, David Oderberg, Edward Feser, John Finnis, F. Russell Hittinger, Pope John Paul II, Saul Kripke, Alvin Plantinga, Elenore Stump, and J.P. Moreland.

⁶⁵ Rawls, "Justice as Fairness: Political not Metaphysical" p394. Emphasis added.

Rawls writes of the primary goods, “these [primary] goods are things citizens need as free and equal persons living a complete life; they are not things it is simply rational to want or desire, or to prefer or even to crave.”⁶⁶ He also tells us that the thin theory of the good is, “taken for granted by any political conception of justice. [...] This idea assumes that human existence and the fulfillment of basic human needs and purposes are good, and that rationality is a basic principle of political and social organization.”⁶⁷ In these passages, Rawls seems to hold that assuming the positive value of fulfilling the desires of citizens poses no threat to justice as fairness.

I have argued that Rawls’ original position thought experiment requires the assumption that fulfilling people’s unspecified rational desires is worthy of choice. Without this assumption, the argument *from* the original position cannot go forward. Furthermore, I have argued that this assumption is metaphysical, in the sense that Rawls employs when he says justice as fairness is to be political and not metaphysical. In my next chapter, I argue that this value ascription is not trivial and that reasonable members of the populace may reject it. As a result, well-meaning individuals willing to abide by fair terms of cooperation may not be able to join in an overlapping consensus establishing Rawls’ two principles of justice. That is to say, making a strong enough assumption so that the argument *from* the original position is sound renders the argument *to* the original position unsound.

⁶⁶ Rawls, *Justice as Fairness: A Restatement*, p58.

⁶⁷ *Ibid.* p141.

CHAPTER FOUR

That an Assumption Regarding Value is Not Neutral

What I have called the ‘trousered ape’ and the ‘urban blockhead’ may be precisely the kind of man they really wish to produce.

—C.S. Lewis, *The Abolition of Man*

In the previous chapter, I argued that the parties behind the veil of ignorance in Rawls’ original position thought experiment have an unannounced commitment to the positive value of the satisfaction of agents’ rational desires. This result, however, by itself has little import. The parties in the original position also seem committed to the idea that more than one person exists – i.e. that solipsism is false. While philosophers continue to wrestle with solipsistic claims, this rejection seems trivial in the context of arranging a system of cooperation between persons. Likewise, Rawls claims that there is a set of facts that are agreed upon and the parties in the original position have access to them. The parties’ knowledge of facts contradicts certain strains of skepticism, including certain metaphysical doctrines, yet this should not seem problematic. All parties inquiring about how to justly organize our political society can reasonably be assumed to have rejected these theses. Some metaphysical claims therefore seem unproblematic, even if the parties in the original position are assumed to rely upon them. The argument *to* the original position does not require zero metaphysical claims, only that there be no *objectionable* metaphysical claims. In this chapter, I argue that the metaphysical commitment argued for in the previous chapter is not a trivial commitment, but rather

prevents the two principles of justice as fairness from being the subject of an overlapping consensus of all reasonable comprehensive doctrines.

I set the stage for this argument by examining Rawls' distinction between the comprehensive doctrine of justice as fairness and the broad family of views that can participate in Rawlsian Political Liberalism. This section will clarify my thesis, showing that it addresses the political conception of justice as fairness, rather than the comprehensive doctrine which may be found in *A Theory of Justice*.¹ Next, I summarize the value ascription I have previously argued constitutes a metaphysical commitment in Rawls' original position thought experiment. This includes discussion of the ways in which Rawls thinks Political Liberalism is neutral.

I then argue that this value assumption is not merely a feature of the original position thought experiment. The value assumption is also expressed in Rawls' doctrine of paternalism and in the two principles of justice as fairness. Therefore, those who have sufficient reason to reject the value assumption because it is incompatible with their comprehensive doctrine also have sufficient reason to reject the two principles of justice as fairness. I do this by showing that some families attempting to raise children must reject Rawls' two principles due to the principle of paternalism Rawls holds to be entailed by these principles. Finally, I argue that such people may be reasonable people holding to reasonable comprehensive doctrines. Therefore, adherents of some reasonable comprehensive doctrines cannot join in an overlapping consensus on the two principles of justice as fairness. The value assumption required to make the argument *from* the

¹ *Ibid.* p xvi-xvii.

original position sound is too strong and makes the argument *to* the original position unsound.

Critiquing Justice as Fairness Versus Critiquing Political Liberalism

In introducing justice as fairness, Rawls writes,

Consider again the idea of social cooperation. How are the fair terms of cooperation to be determined? Are they simply laid down by some outside authority distinct from the persons cooperating? Are they, for example, laid down by God's law? Or are these terms to be recognized by these persons as fair by reference to their knowledge of an independent moral order? For example, are they recognized as required by natural law, or by a realm of values known by rational intuition? Or are these terms established by an undertaking among those persons themselves in the light of what they regard as their reciprocal advantage? Depending on which answer we give, we get a different conception of social cooperation. Justice as fairness recasts the doctrine of the social contract and adopts a form of the last answer: the fair terms of social cooperation are conceived as agreed to by those engaged in it, that is, by free and equal citizens who are born into the society in which they lead their lives.²

Justice as fairness, as set forth in *Theory*, includes this sweeping conclusion that a particular form of contract doctrine correctly fixes the meaning of social cooperation. Yet, if political liberalism is to tolerate a plurality of diverse comprehensive doctrines, political liberalism cannot require all reasonable persons to adhere to such a controversial position. Rawls cannot hope for people from diverse comprehensive doctrines to agree on a comprehensive version of justice as fairness, including meta-ethical claims such as the one cited above. Instead, he hopes that people of diverse comprehensive doctrines can each accept the terms of justice as fairness for reasons based within their diverse comprehensive doctrines. Rawls writes, "since different premises may lead to the same conclusions, we simply suppose that the essential elements of the political conception, its

² Rawls, *Political Liberalism*, p22-23. See also *Political Liberalism* 97, "Justice as Fairness: Political not Metaphysical" p399, and *Justice as Fairness: A Restatement* p14-15.

principles, standards, and ideals, are theorems, as it were, at which the comprehensive doctrines in the consensus intersect or converge.”³

As such, an argument that justice as fairness depends on a controversial metaphysical thesis can only make a very limited impact on Rawls’ position. An example of the limited impact of such criticisms can be seen in the case of Michael Sandel’s critique of *A Theory of Justice*. Sandel argues that Rawls’ argument for employing the original position in *Theory* depends on a Kantian conception of a person.⁴ But Rawls takes this as no threat to political liberalism itself. Rawls writes,

As a device of representation the idea of the original position serves as a means of public reflection and self-clarification. It helps us work out what we now think, once we are able to take a clear and uncluttered view of what justice requires when society is conceived as a scheme of cooperation between free and equal citizens from one generation to the next. The original position serves as a mediating idea by which all our considered convictions, whatever their level of generality – whether they concern fair conditions for situating the parties or reasonable constraints on reasons, or first principles and precepts, or judgments about particular institutions and actions – can be brought to bear on one another. This enables us to establish greater coherence among all our judgments; and with this deeper self-understanding we can attain wider agreement among one another.⁵

And again Rawls warns, “justice as fairness is badly misunderstood if the deliberation of the parties, and the motives we attribute to them, are mistaken for an account of the moral psychology, either of actual persons or of citizens in a well-ordered society.”⁶ All of this business about the original position serves only to help us better understand our settled

³ Rawls, “The Idea of an Overlapping Consensus.” In *Collected Papers* ed. Samuel Freeman. Cambridge, MA: Harvard University Press, 1999. p430.

⁴ Michael J. Sandel, *Liberalism and the Limits of Justice*. Second Edition. Cambridge University Press, 1998. p187.

⁵ Rawls, *Political Liberalism*, p26.

⁶ *Ibid.* p28.

convictions about justice, upon which Rawls expects we will find significant agreement. We do not need to accept that we could possibly find ourselves actually in the original position, we merely play a game and any person is capable of playing the game.⁷ Likewise, it should not matter if we disagree somewhat on the lines of argument embraced by the characters in the roleplay, so long as we agree on the conclusions resulting from the roleplay.⁸ The criticism that the parties in the original position rely upon an ascription of value to the satisfaction of the rational desires of humans may appear to go no deeper than disagreeing on the lines of argument embraced by the characters of the roleplay.

Justice as fairness, as addressed in *A Theory of Justice*, could be a comprehensive doctrine and so any claims required of adherents, while they might provide a reason to reject justice as fairness as a comprehensive doctrine, do not give a reason to reject the political conception. Rawls is clear on this point in *Political Liberalism*, “although the distinction between a political conception of justice and a comprehensive philosophical doctrine is not discussed in *Theory*, once the question is raised, it is clear, I think, that the text regards justice as fairness and utilitarianism as comprehensive, or partially comprehensive, doctrines.”^{9,10} Given the fact of reasonable pluralism, Rawls cannot

⁷ That some citizens will find playing the game morally objectionable is a possibility brought up by Patrick Neal. See “Political Liberalism, Public Reason, and the Citizen of Faith” in eds. Robert P. George and Christopher Wolfe *Natural Law and Public Reason* (Georgetown University Press, 2000).

⁸ See *Political Liberalism* p27.

⁹ Rawls, *Political Liberalism*, p xvi.

¹⁰ In spite of the fact that Rawls seems clear in *Political Liberalism* that Justice as Fairness, as found in *Theory*, was a comprehensive doctrine, I have endeavored to restrain my language to the position that it *may* be a comprehensive doctrine. I do this because, in the introduction to *Restatement*, Rawls tells us that *Theory* is ambiguous in expressing Justice as Fairness as a political conception only, or as a comprehensive doctrine. See *Restatement*, xvii.

insist on a (partially) comprehensive doctrine. Rawls further develops this point in *Restatement*, “these [comprehensive views which can endorse justice as fairness] include religious doctrines that affirm liberty of conscience and support the basic constitutional freedoms, as well as various liberal philosophical doctrines, such as those of Kant and Mill, that likewise do so.”¹¹ What is important to the overall project of political liberalism is the political conception that arises from justice as fairness, rather than agreement on the comprehensive doctrine of justice as fairness. “It suffices to show how a conception of justice with the structure and content of justice as fairness can be understood as political and not metaphysical.”¹²

I intend to show that the criticism raised in the previous chapter extends further than a comprehensive doctrine of justice as fairness. That is, that the ascription of value to the satisfaction of the rational desires of humans required for the soundness of the argument *from* the original position undermines the argument *to* the original position. To this end, we must be clear about exactly what role justice as fairness plays in Rawls’ defense of political liberalism.

Political Liberalism’s Dependence on Justice as Fairness

Rawls explains the task of *Political Liberalism*,

Thus, a main aim of PL is to show that the idea of the well-ordered society in *Theory* may be reformulated so as to take account of the fact of reasonable pluralism. To do this, it transforms the doctrine of justice as fairness as presented in *Theory* into a political conception of justice that applies to the basic structure of society. Transforming justice as fairness into a political conception of justice

¹¹ Rawls, *Justice as Fairness: A Restatement* p33.

¹² Rawls, *Political Liberalism*. p13.

requires reformulating as political conceptions the component ideas that make up the comprehensive doctrine of justice as fairness.¹³

He clarifies the distinction between the political conception and the comprehensive doctrine, “by contrast [to a comprehensive moral doctrine, viz. utilitarianism], a political conception tries to elaborate a reasonable conception for the basic structure alone and involves, so far as possible, no wider commitment to any other doctrine.”¹⁴ For example, *A Theory of Justice* contained a fuller account of the good of social union and a presentation of a system of government in four branches – articles that are not properly part of a political conception of justice.¹⁵ In laying out these four branches of government, *Theory* assumes the acceptability of wealth redistribution, another thesis that cannot be part of a political conception in a pluralistic democratic society. Because these theses go too far for a political conception, Rawls does not continue to defend them in subsequent works.¹⁶ Thus a value ascription contained within justice as fairness, understood as a comprehensive doctrine, poses no problem for political liberalism, whereas a controversial value ascription contained within the political conception described by justice as fairness may pose a problem.

The question for political liberalism then is whether or not people from diverse reasonable comprehensive doctrines can agree on the terms of the political conception laid out by justice as fairness. Rawls states the issue thusly, “the point is that not all

¹³ *Ibid.* p xli.

¹⁴ *Ibid.* p 13.

¹⁵ Rawls, *A Theory of Justice*, p275-276, 520-530.

¹⁶ See also: Gerald Gaus, “Reasonable Pluralism and the Domain of the Political: How the Weaknesses of John Rawls’s Political Liberalism Can be Overcome by a Justificatory Liberalism” in *Inquiry: An Interdisciplinary Journal of Philosophy*, 42:2, (1999) 259-284.

reasonable comprehensive doctrines are liberal comprehensive doctrines; so the question is whether they can still be compatible for the right reasons with a liberal political conception.”¹⁷ Rawls tells us what political liberalism needs to answer this question,

Thus, political liberalism looks for a political conception of justice that we hope can gain the support of an overlapping consensus of reasonable religious, philosophical, and moral doctrines in a society regulated by it. Gaining this support of reasonable doctrines lays the basis for answering our second fundamental question as to how citizens, who remain deeply divided on religious, philosophical and moral doctrines, can still maintain a just and stable democratic society. To this end, it is normally desirable that the comprehensive philosophical and moral views we are wont to use in debating fundamental political issues should give way in public life. Public reason – citizens’ reasoning in the public forum about constitutional essentials and basic questions of justice – is now best guided by a political conception the principles and values of which all citizens can endorse. That political conception is to be, so to speak, political and not metaphysical.¹⁸

Rawls adds the following footnote, “the context here serves to define the phrase, ‘political not metaphysical.’”¹⁹ Again, Rawls gives us an understanding of how the political conception laid out by justice as fairness is to accomplish this; “the aim of justice as fairness, then, is practical: it presents itself as a conception of justice that may be shared by citizens as a basis of a reasoned, informed, and willing political agreement. It expresses their shared and public political reason. But to attain such a shared reason, the conception of justice should be, as far as possible, independent of the opposing and conflicting philosophical and religious doctrines that citizens affirm.”²⁰ Rawls needs a political conception that can be the focus of an overlapping consensus then. In order for

¹⁷ Rawls, *Political Liberalism*, p xxxvii.

¹⁸ *Ibid.* p 10.

¹⁹ *Ibid.* p 10, n10.

²⁰ *Ibid.* p 9.

a political conception to serve this role, it must not conflict with the values of reasonable citizens. Rawls writes,

Political liberalism, then, aims for a political conception of justice as a freestanding view. It offers no specific metaphysical or epistemological doctrine beyond what is implied by the political conception itself. As an account of political values, a freestanding political conception does not deny there being other values that apply, say, to the personal, the familial, and the associational; nor does it say that political values are separate from, or discontinuous with, other values. One aim, as I have said, is to specify the political domain and its conception of justice in such a way that its institutions can gain the support of an overlapping consensus. In this case, citizens themselves, within the exercise of their liberty of thoughts and conscience, and looking to their comprehensive doctrines, view the political conception as derived from, or congruent with, or at least not in conflict with, their other values.²¹

But it is not enough for the political conception to be the subject of an overlapping consensus of the comprehensive doctrines that actually exist in the real world. It must not depend on contingencies regarding which comprehensive doctrines are present in society, lest stability be threatened by other reasonable comprehensive doctrines that might arise which cannot participate in the overlapping consensus. Furthermore, the argument from the original position prevents reliance upon what doctrines are actually present in society by denying this information to the parties while they discuss the basic structure of society. Regarding this requirement, Rawls writes,

For a political conception to avoid being political in the wrong way, it must formulate a free-standing view of the very great (moral) values applying to the political relationship. It must also set out a public basis of justification for free institutions in a manner accessible to public reason. By contrast, a political conception is political in the wrong way when it is framed as a workable compromise between known and existing political interests, or when it looks to particular comprehensive doctrines presently existing in society and then tailors itself to win their allegiance.²²

²¹ *Ibid.* p 10-11.

²² Rawls, *Justice as Fairness: A Restatement*, p188.

But while the political conception of justice laid out by justice as fairness must not conflict with the other values of the reasonable comprehensive doctrines possibly present in society, it does not mean that the political conception of justice does not favor certain comprehensive doctrines over others in terms of indirect promulgation – that is, the relative success that doctrines will have in finding adherents in society.

Justice as fairness is not procedurally neutral. Clearly its principles of justice are substantive and express far more than procedural values, and so do its political conceptions of society and person, which are represented in the original position. As a political conception it aims to be the focus of an overlapping consensus. That is, the view as a whole hopes to articulate a public basis of justification for the basic structure of a constitutional regime working from fundamental intuitive ideas implicit in the public political culture and abstracting from comprehensive religious, philosophical, and moral doctrines. It seeks common ground – or if one prefers, neutral ground – given the fact of pluralism. This common ground is the political conception itself as the focus of an overlapping consensus.²³

In summary, we see that justice as fairness is intended to be one possible subject of an overlapping consensus of reasonable (partially) comprehensive doctrines. In order to serve this role, adherents of all possible reasonable (partially) comprehensive doctrines must be able to agree to be bound by the terms of justice as fairness, provided that the other members of society agree to be bound by the same terms. Rawls does not require them to accept these terms for the reasons specified by a comprehensive doctrine of justice as fairness. He states, “while in a well-ordered society all citizens affirm the same political conception of justice, we do not assume they do so for all the same reasons, all the way down.”²⁴ Rawls needs real world citizens to be able to agree to the two

²³ Rawls, *Political Liberalism*, p 192.

²⁴ Rawls, *Justice as Fairness: A Restatement* p32.

principles of justice as fairness, but he does not need those citizens to be convinced by the original position thought experiment.

I have argued that the parties in the original position thought experiment make a particular value assumption. In the next section, I argue that this commitment is necessary for the two principles of the political conception of justice as fairness themselves, not only for the argument Rawls says those in the original position find compelling. Certain features of a comprehensive doctrine would make this value ascription unacceptable to adherents; I then set out what these features are and argue that these features do not entail that the doctrine is unreasonable.

The Metaphysical Commitment of Justice as Fairness

In the previous chapter, I argued that the representatives in the original position must make value ascriptions in order to arrive at any decision. Without an ascription of value, the parties cannot make a contrastive choice between justice as fairness and other systems as Rawls requires. Without this ascription of value, the argument *from* the original position is unsound. I then examined Rawls' works and noted all the possible sources that may have led to the parties' ascriptions of value. Rawls outlines a particular sense of good, which he calls goodness as rationality or the thin theory of the good. He claims that this sense of good is purely descriptive, yet nevertheless sufficient to enable the parties to carry out their task. On this account, something is good for someone if and only if it helps her satisfy her rational preferences.

Yet this use of the word good, so long as it is purely descriptive, cannot ground ascriptions of choiceworthiness. The parties in the original position would need to perceive the value that their constituents in fact do perceive in their chosen life plans to

explain the parties' motivation to choose principles of justice that advance these life plans. Some evaluative judgment is necessary to motivate the parties to choose the two principles of justice as fairness over another theory as Rawls would have them do and that evaluation requires some ascription of value. At some point, the representatives in the original position accept that furthering the satisfaction of the rational desires of their constituents makes a theory of justice worthy of choice. While it is certainly possible that the satisfaction of some rational desires may turn out to have an objective value that swamps this value, some positive value must be ascribed to the satisfaction of rational desires, qua the satisfaction of rational desires. This value cannot come from the (purely descriptive) thin theory of the good, which Rawls intends to guide the choices of the parties in the original position. Nor can it come from the higher order interests Rawls argues for. The assumption that satisfying individuals' unspecified rational desires is worthy of choice cannot be plausibly included amongst the generally accepted set of facts handed to the representatives in the original position, nor derived from the relationship between the representatives and their constituents. Rawls does not specify a source for this assumption, but the assumption must be made if the parties in the original position are to have a reason to select one theory of justice over another, as the thought experiment requires them to do.

Rawls may have recognized that the parties in the original position ascribe some positive value to fulfilling just rational desires of citizens. He writes, "the first [idea of goodness] is that of goodness as rationality, and in some form it is taken for granted by any political conception of justice. [...] This idea assumes that human existence and the fulfillment of basic human needs and purposes are good, and that rationality is a basic

principle of political and social organization.”²⁵ The same assumption can be seen earlier in Rawls’ work. He writes, “other things equal, they [the parties] prefer a wider to a narrower liberty and opportunity, and a greater rather than a smaller share of wealth and income. *That these things are good seems clear enough.*” (Emphasis mine)²⁶ For these things to motivate the parties in the original position to choose a particular theory of justice, they must be seen by the parties as worthy of choice. If they are seen as choiceworthy because they are primary goods, then they are good because they lead to the fulfillment of unspecified desires. Again Rawls writes, “a fundamental difficulty is that since under reasonable pluralism the religious good of salvation cannot be the common good of all citizens, the political conception must employ, instead of that good, political conceptions such as liberty and equality together with a guarantee of sufficient all-purpose means for citizens to make intelligent and effective use of their freedoms.”²⁷ But this assumes that citizens’ making effective use of their freedoms is a good – presumably because, given that Rawls presumes the citizens to be rational, it leads to the satisfaction of their desires. Clearly, the liberty, opportunity, wealth and income of which Rawls speaks will help the constituents to satisfy their desires, but satisfying the constituents’ desires would then have to be seen as choiceworthy itself.

I make no objection to the first two assumptions that human existence and the fulfillment of basic human *needs* are good. These seem to me as benign as assuming that more than one person exists. I argue, however, that reasonable comprehensive doctrines

²⁵ *Ibid.* p141.

²⁶ Rawls, *A Theory of Justice* p396.

²⁷ Rawls, *Political Liberalism*, pxxxix.

can reject the third assumption that the fulfillment of human *desires* is a good, perhaps in favor of a similar assumption that only some desires purposes are such that the fulfillment of those desires is a good.²⁸ If the two principles are to be political in the right way, however, then followers of reasonable comprehensive doctrines must be able to agree on those values. In the next section, I argue that some possible comprehensive doctrines cannot ascribe positive value to fulfilling unspecified human desires.²⁹

The Value Assumption, Paternalism, and Justice as Fairness

In this section I argue that some comprehensive doctrines will require adherents to reject the value assumption previously discussed and that this rejection will in turn require adherents to reject the two principles which Rawls hopes can attract an overlapping consensus of reasonable comprehensive doctrines. I do this by showing that Rawls believes the two principles entail a particular doctrine of paternalism. According to Rawls' paternalism, one acts for others with unknown desires so as to maximize primary goods for that person.³⁰ I then argue that parents who reject the value assumption will reject Rawlsian paternalism, because acting so as to maximize primary

²⁸ Rawls' phrase "fulfillment of basic human needs and purposes" is ambiguous. Basic human purposes could refer to those basic purposes that humans objectively have as a result of what they are, or it could refer to basic purposes adopted and pursued by an agent. Due to context, I interpret Rawls as intending the latter. The former seems contrary to his general project as Rawls believes we cannot expect agreement on basic purposes of the objective sort due to the burdens of judgment.

²⁹ Rawls is specifically interested, not in fulfilling unspecified desires generally, but fulfilling unspecified desires the fulfillment of which is compatible with fulfilling similar unspecified desires of others – that is, only those unspecified desires that he considers just per the first principle of his theory of justice. This further specification cannot add value to these desires, however. Even if one grants that pursuit of desires that is incompatible with equal liberty for all has *lesser* value, it could be the case that such incompatible pursuits have *negative* value and the compatible pursuits have no value at all. Positive value will not come as the result of a privation.

³⁰ It is infelicitous that the problem cases I wish to raise against Rawls' doctrine of paternalism are cases of parenting. I have embraced the infelicity however in the interest of retaining Rawls' phrasing.

goods for their child (whose desires upon reaching majority are unknown because they are as yet unformed) will interfere with inculcating their understanding of the *relative importance* of character against fulfilling one's desires. In the interests of raising a child with virtuous character, such parents may find themselves forced to choose between setting a proper example for the child and maximizing the primary goods the child will have at majority. Rejecting maximizing primary goods, for instance, in favor of setting the proper example for the child, constitutes a rejection of Rawlsian paternalism - that is, rejecting what Rawls claims to be an entailment of the two principles. As such, these parents cannot participate in an overlapping consensus focusing on the two principles of justice as fairness. My final task in this chapter is showing that a comprehensive doctrine which requires the rejection of the two principles for these reasons might be a reasonable comprehensive doctrine.

Connecting the Two Principles to Rawlsian Paternalism and to the Value Assumption

Rawls considers paternalistic action while still refining the two principles of justice. Rawls raises the topic of paternalism with this, "the problem of paternalism deserves some discussion here, since it has been mentioned in the argument for equal liberty".³¹ Moreover, upon completing his discussion of paternalism, he offers a new statement of the first principle in order to accommodate the results of this discussion.³²

Rawls writes, "we must choose for others as we have reason to believe they would choose for themselves if they were at the age of reason and deciding rationally. [...] Thus the father can say that he would be irresponsible if he were not to guarantee the

³¹ Rawls, *A Theory of Justice* p248.

³² *Ibid.* p250.

rights of his descendants by adopting the principle of equal liberty.”³³ Again, Rawls writes,

Two further stipulations are necessary: paternalistic intervention must be justified by the evident failure or absence of reason and will; and it must be guided by the principles of justice and what is known about the subject’s more permanent aims and preferences, or by the account of primary goods. These restrictions on the initiation and direction of paternalistic measures follow from the assumptions of the original position. The parties want to guarantee the integrity of their person and their final ends and beliefs whatever these are.³⁴

Participants in an overlapping consensus focusing on the two principles of justice are not required to arrive at the two principles through the original position thought experiment – that would amount to requiring comprehensive liberalism. But the parties in the original position choose the two principles in part because, per Rawls, the two principles guarantee Rawlsian paternalism. Rawls concludes his discourse on his of paternalism, “taking the preceding discussion into account, we can reformulate the first principle of justice and conjoin it to the appropriate priority rule.”³⁵ The first principle is put into its final form in order to make certain that the principles of Rawlsian paternalism are part of the basic structure of society.

Put another way, Rawls argues that in order to be acceptable to the parties in the original position, the two principles must entail Rawlsian paternalism. Rawls then tailors the principles to ensure they meet this requirement, indicating that, in Rawls’ understanding, the two principles entail Rawlsian paternalism.

³³ *Ibid.* p209.

³⁴ *Ibid.* p250.

³⁵ *Ibid.* p250.

Therefore, I conclude that if one rejects the Rawlsian doctrine of paternalism, then one ought to also reject the two principles of justice as fairness. Rejecting Rawlsian paternalism is sufficient for rejecting the two principles.

The Rawlsian value assumption is present in this doctrine of paternalism as well. Rawls writes, “as we know less and less about a person, we act for him as we would act for ourselves from the standpoints of the original position.”³⁶ But the actions of the parties in the original position are motivated by the value assumption I have previously discussed. Were it not for this ascription of value to fulfilling the unspecified desires of their constituents, the parties in the original position would not act as Rawls says they do. Indeed, without an ascription of value, they could not act at all. Therefore, when acting according to the Rawlsian doctrine of paternalism, we act in the same way we would if motivated by the assumption that fulfilling undetermined desires is worthy of choice.³⁷ Rejecting the value assumption is sufficient reason to reject Rawlsian paternalism. In turn, rejecting Rawlsian paternalism is sufficient for rejecting the two principles, so one who rejects the value assumption has sufficient reason to reject the two principles of justice as fairness.

In arguing that the Rawlsian principle of paternalism commends the two principles to those situated in the original position, Rawls makes it clear that the eventual

³⁶ *Ibid.* p249.

³⁷ One might object that one can act in this way but for reasons unrelated to the value assumption. This concern is addressed below but can be summarized: acting as if something were true can constitute a tacit endorsement and that tacit endorsement communicates. A parent attempting to prevent a child from believing something to be true undermines her efforts by acting as if that something were true.

acceptability of paternalistic actions to the recipient is a necessary but not sufficient condition for right paternalistic action. He bids us to consider the following situation:

Imagine two persons in full possession of their reason and will who affirm different religious or philosophical beliefs; and suppose that there is some psychological process that will convert each to the other's view, despite the fact that the process is imposed on them against their wishes. In due course, let us suppose, both will come to accept conscientiously their new beliefs. We are still not permitted to submit them to this treatment.³⁸

Even when a person's failure to function licenses paternalistic intervention, forced conversion is not permissible. However, Rawls' discussion omits a critical point - not all cases of instilling a system of religious or philosophical belief are cases of *conversion*. Indeed, the paradigm case of paternalistic action is that of a father caring for his children. Parental care includes instilling character. More than simply transmitting a set of beliefs about the world or the fundamental elements of a comprehensive doctrine, parents are responsible for influencing and shaping the desires that their children will have upon reaching majority, even before the child has significantly developed these desires.

Rawlsian Paternalism – Parenting for Primary Goods

Children's outlook on the world, their character and the things that they desire upon reaching majority are heavily influenced by those responsible for parenting them. Because children's character is formed through their upbringing, Rawls' doctrine of paternalism entails that those raising children must act for those children according to the primary goods doctrine (at least until the children are close enough to maturity). Because the character and desires of a child are not yet determined, those parenting the child cannot possibly know what the child's desires will be upon reaching majority.

³⁸ *Ibid.* p249-250.

Consequently, the Rawlsian doctrine of paternalism dictates that parenting, the paradigm case of paternalistic action, must aim at securing the primary goods for the child upon reaching the age of majority. Rawls writes,

The persons in the original position, however, are prevented from knowing any more about their descendants than they do about themselves, and so in this case too, they must rely upon the theory of primary goods. Thus the father can say that he would be irresponsible if he were not to guarantee the rights of his descendants by adopting the principle of equal liberty. From the perspective of the original position, he must assume that this is what they will come to recognize as for their own good.³⁹

When a child is first conceived and the parents begin their care for their child, their knowledge of that child's desires at majority is as minimal as possible in reality. In fact, they may be parents for some time without knowledge of the child's existence, while still making choices that impact the child's life in myriad ways. Upon discovering the pregnancy, they have some knowledge of the child's general situation in life (race, social class, etc.) to guide them beyond what parties in the original position know about their constituents. But this offers little guidance as to what desires the child will have formed decades later and the parents would thus default to the primary goods approach per Rawls' doctrine of paternalism.

Rawls does discuss educating children under justice as fairness. He writes,

An example may clarify this point: various religious sects oppose the culture of the modern world and wish to lead their common life apart from its foreign influences. A problem now arises about their children's education and the requirements the state can impose. The liberalism of Kant and Mill may lead to requirements designed to foster the values of autonomy and individuality as ideals to govern much if not all of life. But political liberalism has a different aim and requires far less. It will ask that children's education include such things as knowledge of their constitutional and civic rights, so that, for example, they know that liberty of conscience exists in their society and that apostasy is not a legal crime, all this to ensure that their continued religious membership when they

³⁹ *Ibid.* p209.

come of age is not based simply on ignorance of their basic rights or fear of punishment for offenses that are only considered offenses within their religious sect. Their education should also prepare them to be fully cooperating members of society and enable them to be self-supporting; it should also encourage the political virtues so that they want to honor the fair terms of social cooperation in their relations with the rest of society.⁴⁰

What Rawls says about education of children focuses on instilling knowledge of facts. But he does not completely neglect the formation of children's desires as part of the process of education – he wants to make sure that all the children are raised in such a way as to participate in the overlapping consensus on fair terms of social cooperation. Rawls speaks here of encouraging virtues in the child, so he seems to expect that education plays at least some role in the formation of the child's character upon reaching majority. If those responsible for parenting influence the child's character, then the end at which their parenting aims will influence the child's eventual character and desires.

When raising children, those parenting both directly and indirectly influence the eventual desires and worldview of the child. Parents directly teach a great many things to their children, especially basic life skills not found in the curricula of common schools. Further, the parents' actions and the priorities displayed in those actions influence the children. Children notice when parents' actions do not match up with the parents' direct teachings.⁴¹ Parents whose actions on behalf of their children consistently advance the interests in terms of primary goods communicate clearly to their children that obtaining the means to satisfy one's desires is of great importance. Parents unwilling to grant this

⁴⁰ Rawls, *Justice as Fairness: A Restatement*. p156.

⁴¹ This seems common sense. Consider what George Sher and William Bennett say regarding teaching moral values: "What exactly does such teaching [directive teaching of basic moral principles involve? [...]] Of these, perhaps the most important is a teacher or administrator's willingness to demonstrate that he himself endorses certain principles – that he accepts them as guides to his own conduct and expects his students to do likewise." George Sher and William J. Bennett, "Moral Education and Indoctrination" *The Journal of Philosophy* vol. 79 (11), 667.

greater value to fulfilling the unspecified rational desires of people will find themselves unable to communicate this message and will have to reject parenting according to the primary goods model as Rawls requires. The consistent pattern of actions to advance the quantity of primary goods that the child will have upon reaching majority embodies the value assumption first discussed as being employed by the parties in the original position.

Consider a couple undertaking to raise their children so that the children embrace a sacrificial standard of giving.⁴² By sacrificial giving, I mean giving which guarantees that the couple cannot fulfill as many of their desires as they otherwise would be able to. The couple believes, per their comprehensive doctrine, that fulfilling unspecified desires is of no value – that is, they reject the Rawlsian value assumption. They do, however, ascribe value to a particular virtue that is expressed through acts of sacrificial giving. Since they do not know the values their young children will have upon reaching majority, they do not value the fulfillment of those desires. But they do value certain sorts of virtues and thus value certain desires that their children might have upon reaching majority. These values express themselves as the couple seeks to inculcate the right sort of desires in their children, so that the children will have the right desires upon reaching majority.

These parents give sacrificially – motivated not only by the value they place upon sacrificial giving, but also by the value they place on inculcating similar desires and virtue in their children. In so doing, they reduce the quantity of resources they have for

⁴² The reasons for which the couple does this are intentionally ambiguous. The interchangeability of the following options is itself instructive. They may be utilitarian and hold that each ought to give until they derive as much utility from retaining a marginal dollar as the recipient would derive from receiving that marginal dollar. The couple may be Christians who believe that the Bible teaches that giving should demonstrate sacrifice by the giver. With minimal adaptation, their standard of giving might also be fixed at a percentage such that, given their level of income, that percentage happens to be sacrificial for that family.

addressing their other desires, such as being able to afford the highest quality post-secondary education for their children and otherwise providing the children with a larger quantity of primary goods at the age of majority. To maximize the quantity of primary goods which their children will have available to them at the age of majority would require that the family abstain from sacrificial giving. Maximizing their chance to inculcate this virtue in their children requires continuously modeling the behavior to their children – retaining funds to maximize the primary goods available to their children at majority would undermine their efforts to inculcate virtue. The couple rejects the Rawlsian value assumption and must therefore reject the Rawlsian doctrine of paternalism. To accept either of these positions would prevent them from carrying out their life plan to give sacrificially and to care for their children by encouraging their children to have the desires related to the virtues they value.

One objection may arise that this conflict between the couple's efforts to inculcate sacrificial generosity in their children and Rawlsian paternalism is merely an inequality in indirect promulgation. That is, meeting the requirements of Rawlsian paternalism will reduce the parents' chance of succeeding in transmitting the trait, but without preventing them from having and practicing their particular position on sacrificial giving as part of their comprehensive doctrine. Rawls does not claim that justice as fairness is neutral in this way and some doctrines will find such a society more or less conducive to promulgation. But this objection misses the point. If Rawlsian paternalism requires that the parents refrain from engaging in sacrificial giving, then it prevents them from a crucial part of undertaking to train their children in a similar standard of generosity. So long as they view the moral formation of the children (including inculcation of the virtue

related to giving sacrificially) as a greater part of their parental duties than providing primary goods at majority, Rawlsian paternalism conflicts with practicing their comprehensive doctrine with their personal financial resources. Since the couple cannot refrain from doing their best to inculcate virtues related to sacrificial giving, and parenting to maximize primary goods communicates that primary goods are more important than having the right desires, they must reject Rawlsian paternalism.

Moral Formation – Incompatible Comprehensive Doctrines

Justice as fairness is not to be understood as a concatenation of truth claims. In arguing for justice as fairness, Rawls takes pains to avoid making any claims to truth, at least when truth is understood in a robust sense in which an assertion that a proposition is true entails an assertion that reality is a certain way. In *Theory*, Rawls writes, “conceptions of justice are to be ranked by their acceptability to persons [in the original position].”⁴³ He argues that justice as fairness ranks, in this particular sense, above any other theory of justice that has been entertained, rather than arguing that justice as fairness is true in a robust sense. Acceptability, in turn, is understood in terms of being coherent with convictions held in reflective equilibrium, rather than correspondence with reality.⁴⁴ Rawls does not argue that justice as fairness is true, rather he argues that an objectively agreeable procedure, the original position thought experiment, will arrive at the two principles of justice as fairness.⁴⁵ By avoiding making claims about the way the

⁴³ Rawls, *A Theory of Justice* p17.

⁴⁴ *Ibid.* p20

⁴⁵ See Rawls, “Kantian Constructivism in Moral Theory”

world is, justice as fairness avoids directly contradicting the claims of many comprehensive doctrines.

Yet the two principles of justice as fairness require one to make certain practical commitments with one's life and actions. If one cannot embrace these commitments, then one cannot embrace the two principles of justice as fairness.⁴⁶ If one cannot embrace the two principles of justice as fairness, then one cannot participate in an overlapping consensus focusing on the two principles of justice as fairness. Yet all this remains inconsequential for Rawls – he requires an overlapping consensus of *reasonable* comprehensive doctrines focusing on the two principles. Rawls freely admits that there will always be members of society who do not participate in the overlapping consensus, at least not in the right way. Followers of unreasonable comprehensive doctrines may reject the two principles or accept them only as a part of a *modus vivendi*, waiting for a chance to reject the two principles and enforce their own views on the rest of society. Hence, I now argue that people who reject the value assumption do not have to be unreasonable. This task begins with an analysis of what it is for a person to be reasonable and for a comprehensive doctrine to be a reasonable comprehensive doctrine. After that, I argue that these features are compatible with people and comprehensive doctrines that reject the Rawlsian value assumption.

⁴⁶ Here is a parallel argument. 1) Per Rawls, one's actions performed for another must be acceptable to that person when they have (re)gained their faculties. 2) If a child comes to reject the Rawlsian value assumption and does value having objectively correct desires, then the child may not find actions taken on their behalf according to Rawlsian paternalism to be acceptable. 3) Possibly, a child will come to reject the Rawlsian value assumption and instead value having objectively correct desires. Therefore, 4) it is possible that actions guided by Rawlsian paternalism will not meet Rawls' requirements for such actions.

Requirements of Reasonability

Rawls intends the account of reasonable comprehensive doctrines to be at least somewhat inclusive. He writes, “this account of reasonable comprehensive doctrines is deliberately loose. We avoid excluding doctrines as unreasonable without strong grounds based on clear aspects of the reasonable itself. Otherwise, our account runs the danger of being arbitrary and exclusive.”⁴⁷ Comprehensive doctrines are, so to speak, reasonable until proven otherwise. That is, meeting the necessary conditions for being reasonable should also be sufficient for being reasonable. With this in mind, I will undertake to find what conditions are necessary for a person or a comprehensive doctrine to be found unreasonable. This will facilitate my eventual case that having sufficient reason to reject the two principles does not entail meeting the necessary conditions to be found unreasonable.

Rawls lays out two basic aspects of what it is for people to be reasonable. He writes, “the first basic aspect of the reasonable, then, is the willingness to propose fair terms of cooperation and to abide by them provided others do. The second basic aspect, as I review now, is the willingness to recognize the burdens of judgment and to accept their consequences for the use of public reason in directing the legitimate exercise of political power in a constitutional regime.”⁴⁸ On the first basic aspect, Rawls elaborates, “persons are reasonable in one basic aspect when, among equals say, they are ready to propose principles and standards as fair terms of cooperation and to abide by them willingly, given the assurance that others will likewise do so. Those norms they view as

⁴⁷ Rawls, *Political Liberalism* p59.

⁴⁸ *Ibid.* p54.

reasonable for everyone to accept and therefore as justifiable to them; and they are ready to discuss the fair terms others propose.”⁴⁹ Rawls also describes those who are not reasonable, “by contrast, people are unreasonable in the same basic aspect when they plan to engage in cooperative schemes but are unwilling to honor, or even to propose, except as a necessary public pretense, any general principles or standards for specifying fair terms of cooperation. They are ready to violate such terms as suits their interests when circumstances allow.”⁵⁰ Reasonable people then discuss terms of cooperation in good faith, considering the suggestions of others and communicating honestly, rather than using the conversation as a false pretense by which to later obtain a more favorable position and then alter the terms of cooperation. They plan to hew to the terms agreed upon and to find terms genuinely agreeable, rather than using their presently more favorable position to force concessions from others.

Returning to the second basic aspect of reasonable persons, Rawls tells us, “the second basic aspect of our being reasonable is, I have said, our recognizing and being willing to bear the consequences of the burdens of judgment.”⁵¹ Reasonable people then expect disagreement on issues of deep importance and they do not denigrate those who disagree with them because those people have reached different conclusions about weighty and complicated matters. Rather, they respect the complexity of such matters and the diversity of experiences underlying those judgments enough to understand that others, through no blameworthy fault of their own, will inevitably reach different

⁴⁹ *Ibid.* p49.

⁵⁰ *Ibid.* p50.

⁵¹ *Ibid.* p58-59.

conclusions. Reasonable people will not coerce others to act in agreement with their own comprehensive doctrines.

According to Rawls, reasonable comprehensive doctrines have three features:

One is that a reasonable doctrine is an exercise in theoretical reason: it covers the major religious, philosophical, and moral aspects of human life in a more or less consistent and coherent manner. It organizes and characterizes recognized values so that they are compatible with one another and express an intelligible view of the world. Each doctrine will do this in ways that distinguish it from other doctrines, for example, by giving certain values a particular primacy and weight. In singling out which values to count as especially significant and how to balance them when they conflict, a reasonable comprehensive doctrine is also an exercise of practical reason. Both theoretical and practical reason are used together in its formulation. Finally, a third feature is that while a reasonable comprehensive view is not necessarily fixed and unchanging, it normally belongs to, or draws upon, a tradition of thought and doctrine.⁵²

Again, Rawls writes, “recall that reasonable comprehensive doctrines are ones that recognize the burdens of reason [i.e. the burdens of judgment] and accept the fact of pluralism as a condition of life under free democratic institutions, and hence accept freedom of thought and liberty of conscience.”⁵³ Reasonable comprehensive doctrines recognize that unanimity cannot be reached without the use of coercive force because of the burdens of judgment. They also accept that coercive force should not be used to prevent people from believing and practicing their comprehensive doctrine while this freedom is compatible with similar freedom for all.

⁵² *Ibid.* p59.

⁵³ Rawls, “The Domain of the Political and Overlapping Consensus.” In *Collected Papers*, 473–96. Cambridge, MA: Harvard University Press, 1999. p487, n31.

Not My Will But Thine

I have argued that some comprehensive doctrines cannot take part in an overlapping consensus focusing on the two principles of justice as fairness because they cannot endorse the value assumption which Rawls would have the parties in the original position make and which is also expressed in the two principles themselves. But can people who reject this assumption, that satisfying the unspecified just desires of persons has positive value, still satisfy the requirements of being reasonable people? I argue that they can. Rejecting the value assumption renders a comprehensive doctrine unreasonable only if adherents treat the desires of others differently than they treat their own desires. It is possible that adherents regard the satisfaction of their own desires with the same degree of skepticism with which they regard the satisfaction of others' desires. Therefore, it is not the case that rejecting the value assumption renders all such comprehensive doctrines unreasonable. Such reasonable people who reject the value assumption defy Rawls' assumption that everyone has a determinate plan of the good for their life which they desire to be fulfilled. Instead, such a person may form a conception of the good life but will that it be fulfilled only if it is in fact the good life – at the same time willing that they be thwarted (somehow, not necessarily by state power) in their pursuits if they are in fact wrong about their conception of the good life. Any positive value they attach to pursuing and fulfilling their (or anyone else's) desires is conditional upon those desires being the right desires. Such individuals form a conception of the good life and act according to that conception, but simultaneously continue to will that, in the event that they are incorrect, they will be prevented from accomplishing their goals, rather willing that what is best, in objective reality, should occur instead of their intended outcomes.

Rawls tells us that reasonable persons will not insist on truth as they understand it. He writes, “of course, those who do insist on their beliefs also insist that their beliefs alone are true: they impose their beliefs because, they say, their beliefs are true and not because they are their beliefs. But this is a claim that all equally could make; it is also a claim that cannot be made good by anyone to citizens generally.”⁵⁴ At least some adherents of comprehensive doctrines that reject the value assumption, however, do not insist on the truth of their ascriptions of value in dealings with other people. Such adherents take their own fallibility very seriously indeed and refuse to commit to willing the fulfillment of desires based on their own best understanding. Instead, they may simply wish protection against being coerced into acting as if others’ desires are in fact worthy of fulfilling because those who assume unspecified desires are valuable cannot justify this assumption to those who disagree. Those who would coerce them (when they cannot reasonably justify the coercion to those coerced) are the ones whom Rawls declares unreasonable. He states, “it is unreasonable for us to use political power, should we possess it, or share it with others, to repress comprehensive views that are not unreasonable.”⁵⁵

It will also be objected that such persons are not in fact interested in cooperating with others in society. Rawls writes that “reasonable persons [...] are not moved by the general good as such but desire for its own sake a social world in which they, as free and equal, can cooperate with others on terms all can accept.”⁵⁶ Rational agents lack the

⁵⁴ Rawls, *Political Liberalism*, p61.

⁵⁵ *Ibid.* p61.

⁵⁶ *Ibid.* p50.

specific motivation to cooperate according to Rawls, unless they are also reasonable agents.⁵⁷ One could argue that people rejecting the value assumption have no interest in working with others and want to avoid cooperating with other members of society. Hence, these individuals are not reasonable because reasonable people intend to agree upon fair terms of cooperation and to be bound by those terms so long as others likewise are bound by those terms.

In response to this charge, there is no reason to believe that adherents of comprehensive doctrines that reject the value assumption must reject cooperation. Rawls describes the three elements of social cooperation: “cooperation is guided by publicly recognized rules and procedures that those cooperating accept and regard as properly regulating their conduct.”⁵⁸ “Fair terms of cooperation specify an idea of reciprocity: all who are engaged in cooperation and who do their part as the rules and procedure require, are to benefit in an appropriate way as assessed by a suitable benchmark of comparison.”⁵⁹ Those who reject the assumption that the fulfillment of the unspecified desires of rational agents is valuable are not necessarily opposed to such fair terms of cooperation. The idea of reciprocity embodied in the terms that they are inclined to propose may be a different idea of reciprocity than that embodied in the terms that others propose. Perhaps this results from a different understanding of benefitting in an “appropriate way” or choice of a different “suitable benchmark”. In either case, that does not mean that those rejecting the value assumption are not willing to agree to terms that

⁵⁷ *Ibid.* p51.

⁵⁸ *Ibid.* p16.

⁵⁹ *Ibid.* p16.

all can see as fair terms. Rawls immediately continues, “a conception of political justice characterizes the fair terms of cooperation.”⁶⁰ To label a person (or a comprehensive doctrine) as unreasonable on the grounds that they are not interested in “cooperation” then means that the person is labeled as unreasonable simply because they do not agree to the same conception of political justice. If one chooses this line of response to the challenge posed by individuals who reject the Rawlsian value assumption, then for one to be unreasonable only means that that person disagrees with the two principles. An overlapping consensus of reasonable comprehensive doctrines focusing on justice as fairness becomes trivial, meaning nothing more than an overlapping consensus of those doctrines which are part of the overlapping consensus. Proposing different terms of cooperation should not, by itself, render someone unreasonable.

Nor can it be said that these individuals are unreasonable because they are not willing to reciprocate in adhering to rules agreed upon. These individuals, when acting consistently with their own beliefs, do not will the fulfillment of their unspecified desires in something like the original position. Stepping in to their real lives with determinate desires based on perceptions of value, such people still do not will the fulfillment of their desires unconditionally. Rather, even as they pursue the fulfillment of the desire, they only will for the desire to be fulfilled in the case that the desire is genuinely worthy of choice – as a matter of objective fact and not merely the result of a procedure with objectivity. While such individuals are forced to act, they do not have a determinate plan for obtaining their own good as Rawls would have us believe all people do – rather they view it as a determinate plan for obtaining what they believe to be their own good, with

⁶⁰ *Ibid.* p16.

the understanding that this is not the same as having a determinate plan for obtaining their own good. These individuals incorporate fallibilism into their desires for themselves as well as for others. Adherents of the sort of comprehensive doctrine I am interested in here respect the burdens of judgment so deeply that they do not presume their own desires worthy of fulfilling. As such, these individuals cannot be charged with being unwilling to reciprocate in adhering to agreed terms of cooperation – they will others’ desires and their own desires to be fulfilled under exactly the same conditions. As long as these individuals are reasonable, then there may exist at least one reasonable comprehensive doctrine that cannot participate in an overlapping consensus focusing on the two principles of justice as fairness. Identifying those who question the value of fulfilling even their own desires as unreasonable would make Rawls account arbitrary and exclusive.

An Example of Reasonable Rejection of the Value Assumption

What might a case look like in which a family of reasonable persons holds to a reasonable comprehensive doctrine which they believe to be incompatible with the two principles of justice as fairness? I hope to answer this by expanding on an objection to justice as fairness from Patrick Neal in “Political Liberalism, Public Reason, and the Citizen of Faith”.⁶¹ Neal argues that Rawls incorrectly understands the worldview of citizens who believe in, and attempt to submit their lives to, an absolute moral authority. I will attempt to flesh out how such a citizen may reason to the conclusion that they must

⁶¹ Patrick Neal, “Political Liberalism, Public Reason, and the Citizen of Faith.” In *Natural Law and Public Reason*, Robert P. George and Christopher Wolfe, eds. 171–201. Washington, D.C.: Georgetown University Press, 2000.

reject the two principles of justice as fairness and show that such a rejection need not be unreasonable.

Imagine a couple who subscribes to the following argument.

- 1) There is an absolute authority
- 2) If there is an absolute authority, then all value derives from that authority.
- 3) If all value derives from that authority, then the fulfillment of an agent's desires is valuable if and only if those desires align with the absolute authority.
- 4) Therefore, it is more important to have desires that align with the absolute authority than it is to have one's desires fulfilled.

The couple holds to 1) - 3) and thus rejects the Rawlsian value assumption. The couple makes no effort to enshrine any of these beliefs in public policy, however, and is willing to abide by fair terms of cooperation, so long as others agree to abide by those terms as well. They are willing to pay taxes and thereby help fund public institutions in spite of the fact that they believe many of those benefitting from such institutions are pursuing wrong desires. They do these things because they recognize that, due to the burdens of judgment, significant disagreement is expected on many issues.

Recall that there are two basic elements of the reasonable per Rawls. First, people must be willing to abide by fair terms of cooperation provided that others will also abide by them. Secondly, they must have “the willingness to recognize the burdens of judgment and to accept their consequences for the use of public reason in directing the legitimate exercise of political power in a constitutional regime.”⁶² The couple clearly

⁶² Rawls, *Political Liberalism*, p54.

meets these requirements. They are reasonable people adhering to a reasonable comprehensive doctrine.

Now suppose this couple has a child. They believe that it is more important for their child to have the right desires at majority than to maximize the means the child has for satisfying her desires (rejecting the Rawlsian value assumption) and they want their child to see this too. They do not seek to prevent their child from hearing about other views, but they do endeavor to ensure the child knows their view well. To give their view a full-throated voice, they must make it clear to the child that they consider right desires more important than desire fulfillment. To accomplish this, their parenting must emphasize cultivating virtue above the primary goods, so that the child will see this priority expressed in her parents' actions on her behalf. Since they must avoid parenting for primary goods, the couple must reject Rawlsian paternalism. Therefore, they must also reject the two principles of justice as fairness which, per Rawls, would be chosen in the original position because they guarantee Rawlsian paternalism.

One may object that the couple has significant resources with which they are allowed to attempt to inculcate generous character in their children. But the parents are trying to inculcate more than mere generosity, they are attempting to inculcate a thesis about the relative importance of generous character vis-à-vis the importance of having the ability to satisfy one's desires. While Rawls' primary concern is the basic structure of society, his principle of paternalism seems to apply directly to an individual's actions taken on behalf of another, just as a father acts on behalf of his children. Rawls writes that "we act for him as we would act for ourselves from the standpoints of the original

position”, thus showing the personal application of the principle of paternalism.⁶³ Further, the personal impact of Rawlsian paternalism can be seen in the works of contemporary Rawlsian liberals, such as Amy Gutmann and Stephen Macedo.⁶⁴ The impact of Rawlsian paternalism on individuals specific actions creates the potential for conflicts such as previously discussed at the personal level for individuals in society. Some Rawlsians may reject the applicability of Rawlsian paternalism in the context of the family, however, this interpretation seems to dominate current understanding of Rawls’ works.

Summary

I have previously argued that Rawlsian political liberalism assumes that fulfilling people’s unspecified desires is an end worthy of choice. Without such an assumption in the original position, the parties cannot make their choice to agree upon the two principles of justice that Rawls argues they would agree upon. Thus, the argument *from* the original position fails. In this chapter, I have argued that the assumption is not a trivial assumption and that the assumption is present in the political conception of justice as fairness distinguished in Rawls’ later works, not only in a comprehensive doctrine which might be drawn from *Theory*. I then argued that, due to this value assumption, people who adhere to certain comprehensive doctrines will be unable to participate in an overlapping consensus focusing on the two principles of justice as fairness. This runs

⁶³ Rawls, *A Theory of Justice*, p249.

⁶⁴ For the impact some Rawlsians see liberalism having on the formation of children, see Amy Gutmann, *Identity in Democracy* and “Civic Education and Social Diversity”, Amy Gutmann and Dennis Thompson, *Why Deliberative Democracy*, and Stephen Macedo, *Liberal Virtues: Citizenship, Virtue, and Community in Liberal Constitutionalism* and “Liberal Civic Education and Religious Fundamentalism: The Case of God v. John Rawls”.

contrary to what Rawls himself argues, “justice as fairness has the three features of a political conception that should help it to gain the support of a reasonable overlapping consensus. Its requirements are limited to society’s basic structure, its acceptance presupposes no particular comprehensive view, and its fundamental ideas are familiar and drawn from the public political culture.”⁶⁵ The political conception of justice as fairness does not require one specific comprehensive doctrine. But agreeing to the two principles as fair terms of cooperation does require that one hold at least a partially comprehensive doctrine that ascribes positive value to fulfilling the unspecified desires of rational agents. Thus, the argument *to* the original position would be unsound because the value assumption is too strong for the original position to be the correct position from which to determine matters of basic justice.

According to Rawls, the two principles of justice guarantee a certain doctrine of paternalism – in fact the two principles are chosen by the parties in the original position in part because this doctrine of paternalism is entailed by the two principles. This doctrine of paternalism holds that, when the child’s desires at majority are unknown, those acting paternalistically must seek to obtain the greatest possible quantity of primary goods for the child upon reaching majority. Children’s desires at majority are not known, because they are not yet formed. Therefore, Rawlsian paternalism holds that parents must parent with the aim of maximizing the child’s share of primary goods upon reaching majority. This communicates to the child (in a way that words can never fully counteract) that fulfilling his or her own desires is of positive value and that desire fulfillment is of great importance. At least some parents who reject the value assumption

⁶⁵ Rawls, *Justice as Fairness: A Restatement*, p33.

will find themselves unable to make this demonstration to their children out of a perceived duty to inculcate other incompatible character traits instead. These parents cannot agree to the Rawlsian doctrine of paternalism and hence cannot agree to the two principles of justice as fairness.

I finally argued that adherents of such a comprehensive doctrine which rejects the value assumption need not be unreasonable. They are willing to agree to terms of cooperation that can be seen as fair by all – however, they cannot see the terms that Rawls proposes as fair because of the value assumption. This need not result from an unwillingness to reciprocate and abide by whatever fair terms may be agreed upon, however. Some such individuals treat their own desires with the same degree of skepticism – even when they must act according to a perception of the good, they hold that perception lightly and will that their actions succeed only in the case that the actions are based upon accurate perceptions of value and the good. In this, their own desires are treated with the same degree of skepticism that the desires of others are treated and such persons are seen to be willing to reciprocate if a set of terms can be agreed upon by all parties. Finally, I offer the example of a reasonable couple who cannot parent for primary goods because doing so would undermine their efforts to raise children who appreciate the significant possibility that virtue is more important than satisfying their desires.

What does all this mean for political liberalism? Rawls writes,

One reason we form [the idea of a well-ordered society] is that an important question about a conception of justice for a democratic society is whether, and how well, it can serve as the publicly recognized and mutually acknowledged conception of justice when society is viewed as a system of cooperation between free and equal citizens from one generation to the next. A political conception of justice that could not fulfill this

public role must be, it seems, in some way seriously defective. The suitability of a conception of justice for a well-ordered society provides an important criterion for comparing political conceptions of justice.⁶⁶

If my arguments are sound, justice as fairness, as defended by Rawls, cannot fulfill this public role and must be adapted in some way if it is to serve as a publicly recognized and mutually acknowledged conception of justice. In order to do so, the parties in the original position would have to settle for more modest goals in their negotiations. Again Rawls writes, “in justice as fairness, then, the general meaning of the priority of the right is that admissible ideas of the good must fit within its framework as a political conception. Given the fact of pluralism, we must be able to assume: (1) that the ideas used are, or could be, shared by citizens generally regarded as free and equal; and (2) that they do not presuppose any particular fully (or partially) comprehensive doctrine.”⁶⁷ The Rawlsian value assumption violates the second condition and is therefore not an admissible idea of the good.

Rawls has underestimated the diversity of reasonable comprehensive doctrines and thus would have to settle for a less extensive set of regulations for social cooperation. Rawls does consider this possibility; “what if it turns out that the principles of justice as fairness cannot gain the support of reasonable doctrines, so that the case for stability fails? Justice as fairness as we have stated it is then in difficulty. We should have to see whether acceptable changes in the principles of justice would achieve stability; or indeed whether stability could obtain for any democratic conception.”⁶⁸ While the delegates in

⁶⁶ *Ibid.* p9.

⁶⁷ *Ibid.* p141.

⁶⁸ Rawls, *Political Liberalism*, p65-66.

the original position may not be able to secure for their constituents all that Rawls hoped they could secure, they may still hope to at least find terms upon which *all* reasonable persons can agree. If the political conception includes the ascription of positive value to fulfilling the choice of human persons without specification of those desires, then some reasonable citizens cannot agree with the deliberations of the parties in the original position as Rawls presents them. Rawls sums up this problem himself, “if we are to speak of public reason, the knowledge and ways of reasoning – the plain truths now common and available to citizens generally – that ground the parties’ selection of the principles of justice must be accessible to citizens’ common reason. Otherwise, the political conception does not provide a basis of political legitimacy.”⁶⁹

⁶⁹ Rawls, *Justice as Fairness: A Restatement*, p90.

CHAPTER FIVE

Conclusion

I have argued that Rawls' arguments *to* the original position and *from* the original position cannot both be sound. In Chapter Two, I search for a model of civic friendship which could benefit a modern, pluralistic society. In such a society I argue, adherents of very different comprehensive doctrines cannot always recognize others' public actions as expressions of well-wishing. Many real world projects partially address the lack of mutual recognition. Cultural education programs and common schools attempt to provide a common cultural understanding which could facilitate the recognition of good will in a pluralistic culture. But these measures are themselves problematic – for example, required reading lists commonly attract citizens' ire.¹ I propose an account of civic friendship based on Aristotle's model, but replace his use of a robust theory of the good with Rawls' thin theory of the good. This pseudo-Aristotelian model could bind a broader collection of citizens. I conclude, however, that some citizens still cannot participate. Those who ascribe negative value to the fulfillment of unspecified desires of rational agents cannot participate.

Rawls' work assumes that fulfilling the unspecified desires of rational agents has positive value – this assumption then becomes the focus of my inquiry. Yet one should

¹ Consider for example the *Mozert v. Hawkins*, in which Christian families objected to a reading list on the grounds that their religious views were denigrated within some of the works included. Stephen Macedo's "Liberal Civic Education and Religious Fundamentalism: The Case of *God v. John Rawls*" in *The Ethics of Citizenship*, J. Caleb Clanton, ed. (Baylor University Press, 2009) offers significant insight and discussion.

note that those who assign zero value to such desire fulfillment can still participate in civic friendship on my pseudo-Aristotelian model. Pseudo-Aristotelian civic friendship need not break down in the problem cases I lay out for Rawls' project of political liberalism. The pseudo-Aristotelian model is very open and can include citizens who might reject the terms of justice as fairness.

In Chapter Three, I begin investigating this value assumption in Rawls' work. Rawls requires the parties in the original position to act. To explain their action, Rawls' requires a value assumption. The parties, per Rawls, attempt to maximize the ability of their constituents to pursue fulfilling their rational desires. Thus we see that the parties (positively) value the enabling pursuit and fulfillment of their constituents' rational desires. I argue that this assumption is metaphysical – contrary to the Rawlsian slogan that justice as fairness is political and not metaphysical. In doing so, I examine the resources available to the parties in the original position, concluding that none of the information Rawls explicitly attributes to the parties includes this assumption regarding value. Yet this assumption is disputed. If the original position thought experiment imports assumptions not found in the stock of commonly held ideas in society, then the original position is not the correct thought experiment by which to think about matters of basic justice. Thus, his argument to the original position is unsound. But without the value assumption, the argument from the original position cannot move forward as the parties will never act, let alone choose the principles Rawls would have them choose.

In Chapter Four, I argue that because it prevents an overlapping consensus on justice as fairness, this assumption is not a trivial metaphysical thesis. Here my argument depends on Rawls' understanding that the two principles of justice as fairness entail

Rawlsian paternalism. This doctrine of paternalism is incompatible with comprehensive doctrines which hold that inculcating the right desires in children is more important than enabling the children to pursue the fulfillment of the desires they happen to have at majority. Rawlsian paternalism requires parents to maximize the index of primary goods which their children will have at majority. But such parenting undermines the efforts of many parents who adhere to doctrines that require them to live out their comprehensive doctrine before their own family. Parental priorities communicate to children. While Rawls does not claim justice as fairness is neutral in the sense that all comprehensive doctrines will gain similar numbers of adherents, he does claim neutrality in that all reasonable comprehensive doctrines may represent themselves in society with equal assistance from the state. If social institutions enforce Rawlsian paternalism, then such parents would not receive equal assistance to those holding other views.²

A final concern is whether or not such comprehensive doctrines may be reasonable. Individuals are considered reasonable by default. Unless they are unwilling to propose or abide by fair terms of cooperation, or to accept the burdens of judgment and public reason's restriction on the use of coercive state power, people should be considered reasonable in Rawls' sense of the term. In the cases I press, the families in question make no effort to compel anyone through coercive state power, or even to influence state power. They expect disagreement on significant questions and recognize the possibility that their judgments about the nature of the good human life are wrong. Their views are held lightly to the extent that, *ultima facie*, they wish for their desires to

² See Rawls, *Political Liberalism*, p193.

be corrected if said desires differ from the objectively best desires for them to have, rather than wishing to have said desires fulfilled.

Rawls states that the parties in the original position will see Rawlsian paternalism as an entailment of the two principles of justice as fairness. This forms one argument he offers that the parties will choose the two principles, because the parties want to enact Rawlsian paternalism for the benefit of their constituents. But while Rawls holds this position, other Rawlsians may respond by denying that the two principles of justice as fairness in fact entail Rawlsian paternalism.

If the two principles do not entail Rawlsian paternalism, then it remains possible that the parents in the cases I deal with could accept justice as fairness. If no particular doctrine of paternalism can be derived from the two principles, then a more libertarian liberalism could attract an overlapping consensus. Such a liberalism may be able to tolerate both Rawlsians and those who believe that, in parenting, the importance of inculcating right desires trumps the value of enabling pursuit of subjective desires.

BIBLIOGRAPHY

- Anscombe, G.E.M. 2000. *Intention*. Harvard University Press.
- Audi, Robert. 2000. *Religious Commitment and Secular Reason*. Cambridge University Press.
- . 2009. “Religious Convictions and Secular Reasons.” In *The Ethics of Citizenship*, edited by J. Caleb Clanton, 59–92. Baylor University Press.
- Bentley, R. K. 2013. “Civic Friendship and Thin Citizenship.” *Res Publica* 19 (1): 5–19.
- Brewer, Talbot. 2005. “Virtues We Can Share: Friendship and Aristotelian Ethical Theory.” *Ethics* 115 (4): 721–58.
- . 2009. *The Retrieval of Ethics*. Oxford University Press.
- Christman, John. 2015. “Autonomy in Moral and Political Philosophy.” *Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/entries/autonomy-moral/>.
- Cohen, Joshua. 2003. “Deliberation and Democratic Legitimacy.” In *Debates in Contemporary Political Philosophy*, edited by Derek Matravers and Jon Pike. Routledge.
- Cooper, John M. 1999a. “Aristotle on the Forms of Friendship.” In *Reason and Emotion: Essays on Ancient Moral Psychology and Ethical Theory*, 312–36. Princeton, NJ: Princeton University Press.
- . 1999b. “Political Animals and Civic Friendship.” In *Reason and Emotion: Essays on Ancient Moral Psychology and Ethical Theory*, 356–77. Princeton, NJ: Princeton University Press.
- Dreben, Burton. 2003. “On Rawls and Political Liberalism.” In *The Cambridge Companion to Rawls*, edited by Samuel Freeman, 316–46. Cambridge University Press.
- Dworkin, Gerald. 1972. “Paternalism.” *The Monist* 56 (1): 64–84.
- Dworkin, Ronald. 1978. *Taking Rights Seriously*. Harvard University Press.
- Eberle, Christopher J. 2002. *Religious Conviction in Liberal Politics*. Cambridge University Press.

- Finnis, John. 2000. "Abortion, Natural Law, and Public Reason." In *Natural Law and Public Reason*, edited by Robert P. George and Christopher Wolfe, 75–106. Georgetown University Press.
- Gaus, Gerald. 1990. *Value and Justification: The Foundations of Liberal Theory*. Cambridge University Press.
- . 1999. "Reasonable Pluralism and the Domain of the Political: How the Weaknesses of John Rawls's Political Liberalism Can Be Overcome by a Justificatory Liberalism." *Inquiry: An Interdisciplinary Journal of Philosophy* 42 (2): 259–84.
- George, Robert P., and Christopher Wolfe. 2000. "Natural Law and Public Reason." In *Natural Law and Public Reason*, edited by Robert P. George and Christopher Wolfe, 51–74. Georgetown University Press.
- Glass, Ronald David. 2000. "Education and the Ethics of Democratic Citizenship." *Studies in Philosophy and Education* 19 (3): 275–96.
- Gutmann, Amy. 1985. "Communitarian Critics of Liberalism." *Philosophy & Public Affairs* 14 (3): 308–22.
- . 1995. "Civic Education and Social Diversity." *Ethics* 105 (3): 557–579.
- . 2003. *Identity in Democracy*. Princeton University Press.
- Gutmann, Amy, and Dennis Thompson. 1996. *Democracy and Disagreement*. The Belknap Press of Harvard University Press.
- . 2004. *Why Deliberative Democracy?* Princeton University Press.
- Hansot, Elizabeth. 2000. "Civic Friendship: An Aristotelian Perspective." In *Reconstructing the Common Good in Education*, edited by Larry Cuban and Dorothy Shapps, 173–85. Stanford University Press.
- Healy, Mary. 2011. "Civic Friendship." *Studies in Philosophy and Education* 30 (3): 229–40.
- Jenkins, Joyce. 1999. "The Advantages of Civic Friendship." *Journal of Philosophical Research* 24: 459–71.
- Kahane, David. 1999. "Diversity, Solidarity and Civic Friendship." *Journal of Political Philosophy* 7 (3): 267–86.
- Kekes, John. 1984. "Civility and Society." *History of Philosophy Quarterly* 1 (4): 429–43.

- Labaree, David F. 2000. "No Exit: Public Education as an Inescapably Public Good." In *Reconstructing the Common Good in Education*, edited by Larry Cuban and Dorothy Shipps, 110–29. Stanford University Press.
- Larmore, Charles E. 1987. *Patterns of Moral Complexity*. Cambridge University Press.
- Leontsini, Eleni. 2013. "The Motive of Society: Aristotle on Civic Friendship, Justice, and Concord." *Res Publica* 19 (1): 21–35.
- Lewis, C.S. 2001. *The Abolition of Man*. HarperCollins.
- Macedo, Stephen. 1990. *Liberal Virtues: Citizenship, Virtue, and Community in Liberal Constitutionalism*. Oxford University Press.
- . 1995. "Liberal Civic Education and Religious Fundamentalism: The Case of God v. John Rawls?" *Ethics* 105 (3): 468–96.
- Macedo, Stephen, and Patrick J. Wolf, eds. 2004. *Educating Citizens: International Perspectives on Civic Values and School Choice*. Brookings Institution Press.
- MacIntyre, Alasdair. 1989. *Whose Justice? Which Rationality?* University of Notre Dame Press.
- MacIntyre, Alasdair. 1981. *After Virtue*. University of Notre Dame Press.
- Neal, Patrick. 2000. "Political Liberalism, Public Reason, and the Citizen of Faith." In *Natural Law and Public Reason*, 171–201. Washington, D.C.: Georgetown University Press.
- Nozick, Robert. 2013. *Anarchy, State, and Utopia*. Basic Books.
- Nussbaum, Martha C. 2003. "Rawls and Feminism." In *The Cambridge Companion to Rawls*, edited by Samuel Freeman, 488–520. Cambridge University Press.
- Olbrys, Stephen Gencarella. 2006. "Dissoi Logoi, Civiv Friendship, and the Politics of Education." *Communication Education* 55 (4): 353–69.
- Perlstein, Daniel. 2000. "'There Is No Escape... from the Ogre of Indoctrination': George Counts and the Civic Dilemmas of Democratic Educators." In *Reconstructing the Common Good in Education*, edited by Larry Cuban and Dorothy Shipps, 51–67. Stanford University Press.
- Pogge, Thomas W. 1989. *Realizing Rawls*. Cornell University Press.

- Rawls, John. 1971. *Theory of Justice*. Reprint, Cambridge, MA: The Belknap Press of Harvard University Press.
- . 1999a. “Justice as Fairness: Political Not Metaphysical.” In *Collected Papers*, edited by Samuel Freeman, 388–415. Cambridge, MA: Harvard University Press.
- . 1999b. “Kantian Constructivism in Moral Theory.” In *Collected Papers*, edited by Samuel Freeman, 303–58. Cambridge, MA: Harvard University Press.
- . 1999c. *The Law of Peoples*. Harvard University Press.
- . 1999d. “The Domain of the Political and Overlapping Consensus.” In *Collected Papers*, edited by Samuel Freeman, 473–96. Cambridge, MA: Harvard University Press.
- . 1999e. “The Idea of an Overlapping Consensus.” In *Collected Papers*, edited by Samuel Freeman. Cambridge, MA: Harvard University Press.
- . 2000. “Fairness to Goodness.” In *Collected Papers*, edited by Samuel Freeman, 267–85. Harvard University Press.
- . 2003. *Justice As Fairness: A Restatement*. Cambridge, MA: The Belknap Press of Harvard University Press.
- . 2005. *Political Liberalism*. Expanded Edition. New York: Columbia University Press.
- . 2007. *Lectures on the History of Political Philosophy*. Edited by Samuel Freeman. The Belknap Press of Harvard University Press.
- . 2009. *Lectures on the History of Moral Philosophy*. Edited by Barbara Herman. Harvard University Press.
- . 2010. *A Brief Inquiry into the Meaning of Sin and Faith*. Edited by Thomas Nagel and Cohen Joshua. Harvard University Press.
- Reese, William J. 2000. “Public Schools and the Elusive Search for the Common Good.” In *Reconstructing the Common Good in Education*, edited by Larry Cuban and Dorothy Shipps, 13–31. Stanford University Press.
- Sandel, Michael. 1984. “The Procedural Republic and the Unencumbered Self.” *Political Theory: An International Journal of Political Philosophy* 12 (1): 81–96.
- . 1998. *Liberalism and the Limits of Justice*. Second. Cambridge University Press.

- . 2009. “The Public Philosophy of Contemporary Liberalism.” In *The Ethics of Citizenship*, edited by J. Caleb Clanton, 345–64. Baylor University Press.
- Scanlon, T. M. 2003. “Rawls on Justification.” In *The Cambridge Companion to Rawls*, edited by Samuel Freeman, 139–67. Cambridge University Press.
- Scheffler, Samuel. 2003. “Rawls and Utilitarianism.” In *The Cambridge Companion to Rawls*, edited by Samuel Freeman, 426–59. Cambridge University Press.
- Schwarzenbach, Sibyl. 1987. “Rawls and Ownership: The Forgotten Category of Reproductive Labor.” *Journal of Canadian Philosophy* 17 (Supplement 1): 139–67.
- . 1996. “On Civic Friendship.” *Ethics* 107 (1): 97–128.
- . 2005. “Democracy and Friendship.” *Journal of Social Philosophy* 36 (2): 233–54.
- . 2009. *On Civic Friendship: Including Women in the State*. Columbia University Press.
- Schwarzenbach, Sibyl A. 2011. “Fraternity and a Global Difference Principle: A Feminist Critique of Rawls and Pogge.” *International Politics* 48 (1): 28–45.
- Sen, Amartya. 1985. “Well-Being, Agency and Freedom: The Dewey Lectures 1984.” *The Journal of Philosophy* 84 (4): 169–221.
- Sher, George, and William J. Bennett. 1982. “Moral Education and Indoctrination.” *The Journal of Philosophy* 79 (11): 665–77.
- Taylor, Charles. 1989. *Sources of the Self: The Making of the Modern Identity*. Harvard University Press.
- Walzer, Michael. 1990. “The Communitarian Critique of Liberalism.” *Political Theory: An International Journal of Political Philosophy* 18 (1): 6–23.
- . 1991. “The Civil Society Argument.” *Statsvetenskaplig Tidskrift* 94 (1).
- Wolterstorff, Nicholas. 1980. *Educating for Responsible Action*. CSI Publications : Eerdmans.
- . 2009. “The Role of Religion in Decision and Discussion of Political Issues.” In *The Ethics of Citizenship*, edited by J. Caleb Clanton, 143–88. Baylor University Press.
- Woods, Kerri. 2013. “Civic and Cosmopolitan Friendship.” *Res Publica* 19 (1): 81–94.