AUTONOMY AND NEUTRALITY IN LIBERAL POLITICAL THOUGHT

by

Paula L. Watson

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

MASTER OF ARTS IN PHILOSOPHY

APPROVED:

John Christman, Chair

Patrick Croskery

Stephen White

April, 1996 Blacksburg, Virginia

Keywords: Autonomy, Neutrality, Liberalism, Perfectionism, Conceptions of the Good
LD
5655
V855
1936
W387
c.2
AUTONOMY AND NEUTRALITY IN LIBERAL POLITICAL THOUGHT

by

Paula L. Watson

Chair: John Christman

Philosophy

(Abstract)

A central tenet of liberal political philosophy is the claim that the role of the state is to remain neutral among citizens' varying conceptions of the good life. In addition to this tenet, some liberals hold the view that we ought to define freedom as individual autonomy. Moreover, such liberals claim that the state is required to promote and protect freedom conceived of as individual autonomy. Such autonomy-based liberalism is frequently criticized for its efforts to simultaneously maintain both views—that is, a commitment both the value of autonomy and a neutrality condition. Critics often argue that since any substantive normative commitment conflicts with the demands neutrality and commitment to the value of autonomy is a substantive normative commitment, liberals cannot consistently be committed to both a neutrality condition and the value of autonomy. This thesis is concerned with providing the best interpretation of liberalism’s theoretical commitments in a way that is consistent with valuing both autonomy and neutrality.
Acknowledgments

In the process of pursuing my academic interests I have incurred many debts both intellectual and personal. Upon the completion of this thesis it seems appropriate to acknowledge some of those debts. John Christman’s influence with respect to my philosophical endeavors runs much deeper than this present work. I would like to thank John for his contributions in terms of advice, encouragement, patience, and criticism to this thesis. Beyond that I thank John for being an inspiration, since my undergraduate work, to pursue philosophy as a career. I also thank Pat Croskery for being a great teacher and friend. In addition, I thank Pat for his constant willingness to talk with me about my work. I have never walked away from a conversation with Pat without feeling that whatever we were thinking about was now much clearer as a result (unless you count that one Friday afternoon after five, when I asked him to briefly explain Hume’s Treatise). I would also like to thank Stephen White. Stephen, perhaps unknown to him, planted the first seeds of political philosophy in my mind. As an undergraduate I took my first political philosophy class with Stephen and upon completion of the course promptly decided to change my major to philosophy. I also would like to thank Stephen for helping to make Habermas much clearer than I would have been capable of on my own.

On a personal level there are several people whom I would like to pay a debt of gratitude. First, thank you jennifer norford, for being a constant reminder that there is more to doing philosophy than writing words on a piece of paper, and for forcing me to acknowledge my own biases. Second, thank you Mark Russell, for being a constant and unrelenting source of distractions. The time spent talking, biking, hiking, drinking, and just playing in general made working that much easier. Third, I also thank Davina Copsy for helping me to proofread. And lastly, thank you to my family: Mom, Dad, Marti, and Barrett for all your support.
## TABLE OF CONTENTS

Acknowledgments.................................................................................................................. iii

Table of Contents.................................................................................................................... iv

Introduction............................................................................................................................... 1

**Chapter One:** Autonomy and Neutrality Considered......................................................... 6
  I. The Concept of Autonomy................................................................................................. 7
  II. Autonomy as a Precondition of Conceptions of the Good........................................... 10
  III. Haworth and the Subjectivity of Value....................................................................... 18
  IV. Conclusions.................................................................................................................... 22
  V. Endnotes......................................................................................................................... 24

**Chapter Two:** Rawls--The Political Move......................................................................... 27
  I. Elements of Rawls's View............................................................................................... 29
  II. Challenges to Rawls....................................................................................................... 33
  III. Evaluating Political Liberalism.................................................................................... 43
  IV. Endnotes......................................................................................................................... 49

**Chapter Three:** Autonomy and Neutrality Reconsidered............................................... 52
  I. The Concept of Neutrality............................................................................................... 52
  II. Neutral Justification of the Neutrality Condition.......................................................... 57
  III. Autonomy and Neutrality............................................................................................. 65
  IV. Justificatory Neutrality and Autonomy......................................................................... 73
  V. Endnotes......................................................................................................................... 79
Chapter Four:

Perfectionism? .................................................................................................................. 81
I. Perfectionist Morality Applied to Political Institutions ................................................ 84
II. Liberal Justification of the Neutrality Condition ......................................................... 87
III. Conclusions ................................................................................................................ 93
IV. Endnotes ...................................................................................................................... 95

Bibliography ....................................................................................................................... 96
Introduction

In the first place, there is the sphere of what is called personal liberty--a sphere most difficult to define, but the arena of the fiercest strife of passion and the deepest feelings of mankind. At the basis lies liberty of thought--freedom from inquisition into opinions that a man forms in his own mind--the inner citadel where, if anywhere, the individual must rule.

--L.T. Hobhouse¹

Liberalism, as suggested by this quote, is concerned with the protection of personal freedom. As the metaphor of the mind as an “inner citadel” indicates, at least some liberals define personal freedom as self-government--that is, as individual autonomy. Autonomy-based liberalism, then, is concerned with protecting the value of individual freedom conceived of as self-government. In addition to its concern with personal freedom, liberalism is generally characterized as the view that the role of the state is to remain neutral with respect to citizens’ conceptions of the good life. The conjunction of these two tenets--a commitment both to autonomy and to a principle of neutrality--is a focal point for much criticism of an autonomy-based liberalism.

It is argued by some critics of autonomy-based liberalism that analysis of the concept of autonomy reveals a normative conception of persons. This conception of persons is normative insofar as the criteria for self-government include independence as well as certain cognitive capacities, critical reflection, for example. Both the requirement for a kind of independence and the capacity for critical assessment of one’s ends

presuppose a certain vision of the self, that perhaps not all would agree to. For example, some feminists have argued that such criteria reflect particularly “male-biased” conceptions of the person. Moreover, these critics claim that such a normative conception of persons is in tension with the neutrality condition, insofar as any substantive normative commitment conflicts with neutrality. Based upon this purported tension, various critics of liberalism conclude that liberalism is either incoherent or is a perfectionist doctrine and thus not neutral in the appropriate way.

The purpose of this thesis is to assess various strategies for avoiding any tension between a commitment to the value of autonomy and the inclusion of a neutrality condition within liberalism’s theoretical commitments. In doing so, I would like to determine whether a commitment to the value of autonomy is indeed in tension with a neutrality principle, and, if so, what implications this has for an autonomy-based liberalism.

One such strategy for avoiding any tension between autonomy and neutrality is to claim that to have a conception of the good one must be autonomous. This strategy purportedly avoids any tension by making autonomy a precondition of having a conception of the good at all. In other words, if one is non-autonomous, then one lacks a conception of the good, according to this view. It follows then that an autonomy-based liberalism is neutral among all possible conceptions of the good, because only having autonomously held ends counts as having a conception of the good. I will consider this strategy in the First Chapter, in which I argue that such a strategy will not work. This
strategy for avoiding the tension between a commitment to autonomy and a neutrality condition fails, I argue, because it rules out the possibility of non-autonomous conceptions of the good and, yet, some such conceptions ought to count as legitimate in appraisals of people’s lives.

The second strategy I consider is reflected in the work of John Rawls, and attempts to avoid the tension between autonomy and neutrality by conceiving of autonomy as part of a purely “political” conception of the person. This politicalization of the concept of autonomy purportedly avoids the tension with neutrality because one need not be committed to the value of autonomy in all spheres of life to adopt this as a principle. In other words, autonomy is not a comprehensive value. Rather, autonomy is a value in the political realm, where a “political value” is distinguished from a “comprehensive value” in terms of the scope of the value. In response to Rawls’s attempt to construct a merely political conception of autonomy, I draw on recent critiques of this view put forth by those working on a communicative approach to ethics, in particular, Jurgen Habermas. Such critiques suggest that Rawls’s account reduces to asserting that autonomy is indeed a comprehensive value. In other words, I argue that, as it turns out, Rawls’s view essentially does the same work as asserting the value of autonomy as a comprehensive moral value. Insofar as this is accurate the political move does not help to reconcile the tension between autonomy and neutrality.

Based on the limitations and concerns raised about the strategies discussed in Chapters One and Two, I return, in the Third Chapter, to reconsider the compatibility of
the value of autonomy, as a comprehensive value, with a neutrality condition. In this chapter, I focus on the concept of neutrality, and what is entailed therein, in an attempt to be clear about what the neutrality condition truly demands of the liberal state. I argue that, in part, the apparent tension between a commitment to the value of autonomy and the neutrality condition is due to a failure to understand the requirements of such a condition. I offer an analysis of the concept of neutrality, borrowed from Jeremy Waldron, and argue that as liberals we are not required to be neutral among all possible conceptions of the good.

Moreover, I argue that to interpret the neutrality condition as demanding neutrality among all possible conceptions of the good is equivalent to demanding a neutral justification of the neutrality condition, a view which some liberals defend. After careful consideration of this interpretation of the neutrality condition, I conclude that any liberalism which demands a neutral justification of the neutrality condition is incoherent. I then suggest a way in which we might understand the neutrality condition and how this is consistent with the state valuing and promoting autonomy.

The Fourth Chapter is a reflection on the conclusions presented in the Third Chapter. My aim here is to assess the view of liberalism offered in the Third Chapter in light of recent critiques that such autonomy-based liberalism is a perfectionist doctrine. In other words, if it turns out that the view of liberalism presented here reduces to a perfectionist view—that is, a view which articulates a teleological conception of human flourishing and thinks it is the state’s responsibility to promote that conception—then
liberalism fails to be neutral in the appropriate way.

This chapter accomplishes three things. First, I articulate the criteria of a perfectionist view. Second, I consider a perfectionist morality as applied to political institutions or arrangements. Thirdly, I consider a perfectionist justification of non-interference on behalf of the state in its citizens’ lives in order to contrast such a justification with a liberal justification of non-interference on behalf of the state (what I have been referring to as the neutrality condition). Based on this I suggest that what distinguishes autonomy-based liberalisms from perfectionist views is the contrasting kind of justification offered to ground a commitment to a neutrality condition (or a principle of non-interference).
Chapter 1

Autonomy and Neutrality

Autonomy-based liberalism has faced severe criticism; critics claim that a commitment to the value of autonomy is in tension with the central tenet of liberal political philosophy, namely, the neutrality condition of liberal political principles.\(^1\) The neutrality condition is generally understood as the requirement that the state must remain neutral among citizens’ conceptions of the good life. This definition of the neutrality condition is, of course, vague. I intend to clarify what is entailed by such a condition in the Third Chapter, where I undertake an analysis of the concept of neutrality. At this point, though, one should read such a condition as the requirement that the state remain neutral among all possible conceptions of the good. Since this is the interpretation most critics offer when critiquing the compatibility of autonomy and neutrality, it is the interpretation I will use for now.\(^2\)

At least some liberals seek to avoid any tension between valuing autonomy and a commitment to a neutrality condition by claiming that a commitment to the value of autonomy is a presupposition of every conception of the good. This claim consists of two parts. The first part is the claim that to count as the possessor of a conception of the good one must be autonomous with respect to that conception. And second, in virtue of that fact (to have a conception of the good one must be autonomous) one is committed to the value of autonomy, whether or not one (a particular person or the state) explicitly
recognizes that commitment. The first part of the claim is that one must be autonomous with respect to some end, goal, or project in order for that end, goal, or project to amount to a conception of the good. The structure of the second part of the claim is that in virtue of the fact that we value our conceptions of the good, and since only autonomous ends count as conceptions of the good, we are committed to the value of autonomy. In order to be in a position to evaluate the compatibility of these claims about the nature of autonomy, conceptions of the good, and the neutrality condition we need first look at the concept of autonomy.

I. The Concept of Autonomy

The concept of autonomy has been cashed out in many different ways. There is no consensus as to what the concept entails. This lack of agreement, however, does not prevent one from identifying certain central features present in any account of autonomy. There are, I think, three such features, all connected with an individual’s ability to be self-governing. To be self-governing requires that in some sense one’s goals, desires, beliefs, ends, etc. must be genuinely one’s own. This is the authenticity condition. Two additional conditions grow out of this requirement. First, to have authentic ends one must be able to assess or make an evaluation of those ends; this is the requirement that to be autonomous one must be capable of critical reflection. Second, one must exercise procedural independence in the formation of one’s desires. Procedural independence requires both that one’s desires not be formed in non-autonomous ways (e.g., as a result
of manipulation or coercion) and that one’s reflection and evaluation of those ends not be influenced by autonomy inhibiting factors. Of particular interest to the argument I will make below are the authenticity condition and the requirement for critical reflection. Let me say a bit more about each of these conditions and their connection with the ascription of value to one’s ends.

Critical reflection is the requirement that to be capable of self-government one must be able to assess or make an evaluation of one’s ends. In other words, one must have an additional psychological attitude toward one’s desires other than merely wanting them to count as autonomous. Some have characterized the notion in terms of wanting to want the end in question. For example, Gerald Dworkin and Harry Frankfurt offer what has been called a “hierarchical” account of autonomy, which basically states that to be autonomous one must reflect on one’s desires and identify with them.5 The hierarchical aspect is clearly seen through the distinction of higher and lower-order desires. The view is as follows: To be autonomous one must have a higher-order desire that has as its object a lower-order desire and the content of the higher-order desire must be approval of the lower-order desire.6 Not all accounts of autonomy are hierarchical, though.

Gary Watson rejects a hierarchical view, but nevertheless counts critical reflection as a necessary condition for autonomy. Watson identifies two aspects of agency. He claims that agents have a valuational system and a motivational system. The motivational system is simply the “set of considerations which move one to action.”7 The valuational system of an agent is made up of the agent’s values which Watson defines as, “those
principles and ends which he [any agent]--in a cool and non-self-deceptive moment--articulates as definitive of the good." So although Watson requires that agents reflect on their beliefs and desires for those ends to be autonomous, he does not do this in terms of higher and lower-order desires.

Lawrence Haworth cashes out the requirement for reflection or evaluation in terms of critical competence. An agent is competent when he or she "strives to be an effective producer of intended effects." An agent becomes critically competent when she or he makes an evaluation of her or his desires and ends--that is, one's competence becomes critical when, in addition to pursuing particular ends, one takes an evaluative stance toward those ends.

The requirement that one reflect on one's ends is meant to capture the idea that to value something one must judge it as good, where the notion of what is good is distinguished from what is simply desired. Notice, though, that the concept of autonomy requires that the conditions under which reflection occurs are of a certain sort, namely, conditions which allow for authentic reflection. The authenticity condition is meant to capture the idea that for one's ends to be "truly" one's own those ends must be genuinely held. It is from this claim that some autonomy-theorists go on and make the further normative claim that for one's ends to merit value, one must be autonomous with respect to those ends. Haworth argues not only that autonomy is a value, but that one's ends, desires, goals, etc. are worthy of value only in virtue of the fact that they are autonomously held. It is not the case that one's ends, desires, goals have value
simpliciter. Rather, one must stand in a certain kind of relationship to one’s ends for those ends to have value; one must hold those ends autonomously on Haworth’s view.

II. Autonomy as a Precondition of Conceptions of the Good

The familiar philosophical distinction between descriptive and normative claims is germane to this discussion of the concept of autonomy, insofar as the concept may be used descriptively or evaluatively. On the one hand, one may offer a description of the conditions one must meet to count as autonomous. For example, the criteria mentioned in Section I, the capacity for critical reflection, procedural independence, etc. serve as a description of what is required for a person to be self-governing. There is nothing inherently evaluative in this description itself.

On the other hand, one may make the further normative claim that only those ends that meet the conditions required by the concept of autonomy are valuable. These are two fundamentally different claims, and the second does not follow from the first. In other words, it does not follow from any description of what it means to be autonomous that only autonomously held ends are valuable. One must give additional reasons beyond the description of the conditions required for self-government in order to support the claim that only those ends relative to which one is self-governing are valuable. Insofar as the claim that autonomy is a precondition of having a conception of the good serves to distinguish “true” conceptions of the good from non-conceptions of the good, it is an evaluative claim.
Recall that earlier I identified two components of the claim that to count as a conception of the good one must be autonomous relative to that conception. The first component of the claim is that one must be autonomous with respect to one’s ends, goals, and projects for those ends to count as a conception of the good. It might seem as though one is making a descriptive claim to say that to count as having a conception of the good one must be autonomous relative to that conception. But, I think, if we reflect on the notion “conception of the good,” we quickly realize that to refer to some end, project, or goal as a “conception of the good” is to make an evaluation of that end, goal, or project. Moreover, it is to make a positive evaluation of that end, goal, or project. This is particularly evident when one considers that, all else being equal, to say some end is a good (or part of a conception of the good) is to say that at least one person has an interest (the person whose end it is) in achieving that end.

The structure of the claim that to have a conception of the good one must be autonomous relative to that conception is as follows: First, there is the implicit recognition that people have ends, goals, and projects. Second, defenders of the view in question claim that we can distinguish those ends that are conceptions of the good life from those ends that are not in virtue of the relation the person has to those ends. In other words, if one is autonomous (i.e., capable of critically reflecting on those ends, acquired them in the appropriate way, and holds them authentically), then those ends amount to a conception of the good. Implicit in this claim is the claim that we ought to promote those conceptions of the good. To say some end is a conception of the good is to say, all else
being equal, that is a value. If something is a value, all else being equal, we have a reason to promote it. In other words, such ends (i.e., those ends that are conceptions of the good) are valuable, and thus we ought to promote them over non-autonomous ends (non-values).

The second part of the claim that autonomy is a presupposition of any particular conception of the good is that in virtue of that fact we are committed to the value of autonomy. This is an explicitly normative claim. It is normative in two respects: First, it identifies autonomy as a value. Second, the claim is that in virtue of the fact that autonomy is a presupposition of every conception of the good, which is a normative claim, we are committed to the value of autonomy. In other words, the claim is if x (autonomy) confers value on y (conceptions of the good) and we are committed to the value of y’s, we are then committed to the value of x. This brings out nicely the point that defenders of this view rely on the implicit assumption that conceptions of the good are valuable. They think it obvious that we are committed to the value of our conceptions of the good and this is what commits us (or the state) to the value of autonomy.

Two examples of contemporary liberals who claim that autonomy is a precondition for possessing a conception of the good are Jeremy Waldron and Will Kymlicka. Waldron holds the view that to have a conception of the good one must be autonomous with respect to that conception. In his *Liberal Rights*, he writes:

[A conception of the good] is not just the exercise of scheduling our activities so
as to maximize over the course of our lives the satisfaction of the preferences we happen to find ourselves afflicted with. The development of a conception of the good is seen as a more reflective business than that. It involves a process whereby the individual stands back and distances himself, from time to time, from his own desires, and determines autonomously whether these are the sorts of desires he wants to be motivated by. In choosing his motivations, rather than regarding them as mere afflictions, the individual associates the business of binding his life into a unity with a process of evaluation: each tries to determine a basis for his action that will be good by his own lights.11

Waldron is arguing, here, that to count as a conception of the good, one’s ends must be of a certain sort--namely, autonomous. In other words, one must stand in a certain relation to one’s ends for them to merit the status of a conception of the good--that is, one must reflect on those ends and autonomously choose them. Now Waldron does not provide us with an account of what it is to be autonomous, but he does hint at such a conception requiring reflection, evaluation, and a kind of unity among one’s desires; this is consistent with the view of autonomy presented in the first section of this paper.

Will Kymlicka holds a similar view to the one currently being considered. Kymlicka characterizes the claim that to count as possessing a conception of the good one must be autonomous relative to that conception in the following way:

Our projects are the most important things in our lives, but since our lives have to be led from the inside, in accordance with our beliefs about value, we should have the freedom to form, revise, and act on our plans of life. Freedom of choice, then, isn’t pursued for its own sake, but as a precondition for pursuing those projects and practices that are valued for their own sake.12

One should read this passage with an emphasis both on the phrase “led from the inside” and “precondition for pursuing...” Essentially, what Kymlicka is arguing is that it is in
virtue of being acquired in the “appropriate” way—that is, autonomously—one’s ends are valuable. As he puts it, the freedom to form, revise, and choose one’s plans and projects is a condition of those projects and plans becoming worthy of value; this represents the claim that autonomy confers value on conceptions of the good and in virtue of the fact that we value our conceptions of the good we are committed to the value of autonomy.

Waldron and Kymlicka share the view that to possess a conception of the good, one must be autonomous with respect to that conception. It follows from that claim, then, that if one is non-autonomous with respect to some end, goal, or project, that end, goal, or project does not count as a conception of the good. In Waldron’s words, it is a mere affliction or occurrent desire and such desires do not amount to a conception of the good.

Lawrence Haworth spells out an argument to the effect that non-autonomous ends lack the status of conceptions of the good. He refers to his more general view that all values are derivative from the value of autonomy as “autonomism.” Autonomism is “the view that personal autonomy is a fundamental value that conditions such other values as freedom [liberty] and satisfaction, so that the latter depend for whatever (my emphasis) value they have on the presence of the former.” Haworth applies this view to the value of conceptions of the good as well. So in order for any particular end, goal, desire, or project to merit value, on this view, it must be held autonomously. Defenders of the view that to count as a conception of the good one must be autonomous with respect to that conception are making a normative claim that itself must be justified.

The justification that Haworth offers for the claim that for some end to count as a
conception of the good one must be autonomous relative to that conception is for that end to “truly” count as that person’s end it must be genuinely held—that is, the end in question must be authentically held. Haworth writes: “And the notion that a person’s conception of the good should be autonomously arrived at simply expresses the importance of that conception’s genuinely being one’s own.”14 In other words, for any end to count as “my end” it must be autonomously held. In addition to merely describing the conditions of autonomy, Haworth makes the further normative claim that one must stand in a particular relation to one’s ends for those end’s to have value; one’s ends must be authentic. To see this, consider the emphasis on the phrase “genuinely one’s own.” The claim is that for some end to be attributable to a particular individual it must be authentic—that is, capable of being critically reflected upon and acquired in the appropriate way. Otherwise, if an end is not authentic it is not attributable to that individual. The end is question is not truly or genuinely that person’s.

This view of autonomy conferring value on other ends is compatible with the neutrality condition insofar as if one accepts the claim, then a commitment to the value of autonomy is automatically neutral among all possible conceptions of the good because conceptions of the good are defined in terms of autonomously held ends. So the “autonomism” view is neutral among all conceptions of the good insofar as only autonomously held ends count as conceptions of the good. But lets consider this claim, namely, that only autonomous ends have value or count as conceptions of the good. I think the best way to do so is to consider a position that rejects such a view.
Consider a utilitarian who holds the view that utility should be defined as pleasure and disutility as pain.\textsuperscript{15} Such a utilitarian is not at all concerned with the relationship one has to one’s ends. The mere having of desires that produce pleasure is enough to merit attributing value to the satisfaction of that desire. As I pointed out above, Haworth’s response to this position is to point to the phrase “one’s ends” and claim that in order for an end to count as “one’s” end, the end in question must be authentic.\textsuperscript{16} In other words, Haworth must claim that only autonomous desires or ends count as a person’s desires.

Consider the following: If I am manipulated into desiring, say, chocolate ice-cream, whether or not I hold this desire autonomously, I am gratified when the desire is met. The desire is mine in the sense that it is attributable to my body. Claiming that a desire must genuinely be one’s own in order for that desire to count as worthy of value rules out the possibility that the mere satisfaction of desires has value. Now as I pointed out above, Haworth will point to the phrase “my desire” and claim that the cashing out of the notion “my” reveals that a person must have genuine desires for them to count as that person’s.\textsuperscript{17} But it is not clear what this could mean without obviously begging the question. The desire is mine in the sense attributable to my body and to suggest that it is not mine because it is non-autonomous simply assumes the thing that needs to be shown.

In addition to ruling out these sorts of manipulated desires as legitimate conceptions of the good, some critics argue that the view that to possess a conception of the good one must be autonomous with respect to that conception rules out certain sorts of relational desires. For example, some feminist critics argue that a commitment to
autonomy is biased toward individual conceptions of the good, and thus non-neutral in the sense that autonomy is not a value esteemed by all possible conceptions of the good. These critics, then, point to certain sorts of desires that are relational—that is, intimately connected with others and their desires—and argue that such desires ought to count as conceptions of the good. The issue turns on whether one thinks it appropriate to rule out the sorts of desires mentioned here as legitimate conceptions of the good. The view considered here does so by definition, and gets neutrality, interpreted in doing so.

As considered thus far the view that to have a conception of the good one must be autonomous with respect to that conception may succeed in being neutral, but only insofar as it defines conceptions of the good as those ends that are autonomously held. I suggest that we have reason to think that it is not neutral among all possible conceptions of the good insofar as some ends, ruled out by definition, ought to count as legitimate conceptions of the good. Beyond these considerations, though, the view in question still needs to provide us with a reason for thinking that neutrality is demanded by the state, even if it is only among autonomously held conceptions of the good. Haworth attempts to provide a justification of the neutrality condition based on appeals to concerns for the subjectivity of value, while at the same time maintaining the view that an individual must be autonomous with respect to her ends for those ends to count as a conception of the good. I consider the plausibility of this justification in the next section.
III. Haworth and the Subjectivity of Value

In this section, I am concerned to rule out one seemingly attractive justification of a neutrality condition for such autonomy-based versions of liberalism. Such a justification includes an appeal to pluralism, where pluralism is interpreted as a the rejection of objective values. Though Haworth does not develop a full theory of liberalism in the argument I consider here, he does provide an argument from pluralism (conceived of as subjectivity of value) to the normative conclusion that “no one’s conception of the good is preferable to anyone else’s,” and on the basis of this infers that one is not justified in imposing a conception of the good on anyone else.\(^\text{19}\) This parallels a liberal argument for the justification of a neutrality condition, where it is inferred that if all values are subjective, then there is no reason for promoting any particular conception of the good over another; therefore, the state ought to remain neutral among citizens’ conceptions of the good.

However, in light of a commitment to the value of autonomy, an appeal to the fact that there are a plurality of subjective values is inadequate grounds to justify a commitment to a neutrality condition. This is so because the claim that pluralism entails the subjectivity of value is in tension with the view that one’s ends, goals, and desires need to be authentic to merit value. Once this tension is revealed, we have grounds for ruling out one justification for liberal neutrality—namely, the view that we can ground a commitment to neutrality on an argument from pluralism, where pluralism is interpreted as the claim that values are subjective.
My concern here is with the “argument from the subjectivity of value” which relies, as Haworth claims, on pluralism or “the rejection of ideal values.” Haworth uses this argument to conclude that since “no one’s conception of the good is preferable to anyone else’s,” one is not justified in imposing one’s conception of the good on anyone else.\textsuperscript{20} I will return to this below; for now let us focus on the details of Haworth’s argument.

The argument can be recapitulated as follows:

A. The value of a particular end, desire, or goal is solely a function of an individual’s valuing it. In other words, the attribution of value is subjective.
B. It is the case that there are a plurality of subjective values.
C. Since there are a plurality of subjective values, there are no ideal values.
D. Since there are no ideal values, no one’s conception of the good--ends, goals, desires--is preferable to anyone else’s.
E. If no one’s conception of the good is preferable to another’s (there are no ideal values), then no one is justified in imposing his or her conception of the good on another.
F. Therefore, we ought not impose our conception of the good on anyone else; we ought to respect others’ autonomy, their ability to realize what counts as a value.

To show the parallel argument that a liberal might use as a justification for the neutrality condition, one must make the following additional inference:

G. Since no citizen ought to impose his or her conception of the good on another, the state ought to remain neutral among citizens’ conceptions of the good in order that it not privilege, promote, or impose any particular conception of the good.

Let’s focus on the first premise (A) in conjunction with the account of autonomy presented in Section I and Section II. The first premise (A) claims that the answer to a
particular question of whether a particular end or desire has value is up to the person who's end or desire it is--i.e., values are subjective. In addition to merely describing the conditions of autonomy, Haworth makes the further normative claim that one must stand in a particular relation to one's ends for those ends to have value; one's ends must be authentic. Haworth maintains that this claim is consistent with the subjectivity of value.

Recall the example of the person manipulated into desiring chocolate ice-cream mentioned above. If the attribution of value were a subjective matter, then if I say it is a good thing that my desire to eat chocolate ice-cream is satisfied, then the fact that it is non-autonomously held is irrelevant. Now the defender of this position will offer reasons for thinking that only autonomous desires have worth, but this is just what it is to conceive of the attribution of value as an objective matter. Because the defender of the view that value attaches only to autonomously held ends is offering, in effect, standards for when to assign value, she or he rejects the general subjectivity of value. So to suggest that some desire or end that I might hold is not worthy of value because it is not held in the appropriate way (i.e., autonomously) is at the same time to be committed to the view that the attribution of value is an objective matter. In other words, Haworth offers an objective conception of value. So his reason for being committed to a neutrality condition (in light of his commitment to the value of autonomy) cannot be based on an argument from the subjectivity of value.

To see this let us now turn our attention to (A) and (B) in the argument above. As the argument is currently formulated, these reflect Haworth's view of pluralism. The
second premise asserts that pluralism is true—that a multiplicity of subjective values exists. From this it is supposed to follow both that there are no ideal values (C) and “no one’s conception of the good is preferable to anyone else’s” (D), but given what we have said about autonomy this conclusion cannot remain as it now stands in conjunction with (A) (the subjectivity of value claim). This is so because the account of autonomy presented here provides reason for thinking that autonomous preferences are more valuable than non-autonomous preferences.

So an appeal to the general subjectivity of value is insufficient to ground a commitment to a neutrality condition in light of a commitment to the value of autonomy. Faced with this tension, and assuming one maintains a commitment to the value of autonomy, there seem to me two possible avenues: First, one may abandon the general subjectivity of value and maintain the view that only autonomously held ends count as conceptions of the good; this is consistent with a neutrality condition interpreted as the claim that the state must be neutral among all possible conceptions of the good. But, such an account rules out, by definition, non-autonomous preferences as possible conceptions of the good. I suggested that this is undesirable both because it seems plausible that such preferences ought to count as legitimate conceptions of the good and because defining conceptions of the good in terms of autonomy may beg important questions. Secondly, then, we could allow that non-autonomous preferences count as conceptions of the good, maintain a commitment to the value of autonomy, and reconsider whether the neutrality condition demands that the state remain neutral among all possible conceptions of the
good. I opt for the latter strategy, and address just these sorts of concerns Chapter Three.

IV. Conclusions

In sum, the purpose of this chapter is to raise some concerns and point to limitations with regard to the strategy that seeks to avoid any tension between autonomy and neutrality by claiming that autonomy is a precondition of all conceptions of the good. The weakness of this strategy is that its defenders must claim that all non-autonomous ends fail to count as conceptions of the good. In other words, if one does not stand in the appropriate (i.e., autonomous) relation to one's ends, then one lacks a conception of the good, on this view. However, as we saw, in order for this claim to have force, one must provide some justification for the claim. The justification considered here is the claim that to count as one's own ends those ends must be genuine—that is, capable of being critically reflected upon, and acquired in the appropriate way. This justification seems inadequate insofar as it may beg important questions with regard to the value of various ends that are not autonomously held.

The intuition that defenders of this view are trying to capture, I think, is the idea that there is something valuable about autonomous conceptions of the good relative to non-autonomous conceptions of the good, but I suggest that one ought provide a reason to support that claim, rather than merely define conceptions of the good in such a way as to exclude non-autonomous ends. Moreover, I suggest that, in part, the problem in thinking that one must define conceptions of the good only as those ends that are autonomously
held is due to thinking that the neutrality condition demands neutrality among all possible conceptions of the good. I will, in the Third Chapter, consider the concept of neutrality in detail with the goal of getting clear about what is demanded of the liberal state when we say it ought to be neutral among citizens’ conceptions of the good life. Before doing so, I consider another strategy for avoiding any tension between autonomy and neutrality; this strategy involves seeing autonomy as a political value—not as a comprehensive value—and is found in the work of John Rawls.
Notes to Chapter One:


2. I infer that this is the notion of neutrality critics have in mind. I do so because the structure of such critiques is to point to conceptions of the good that are left out by the concept of autonomy, and on this basis claim that a commitment to the value of autonomy is non-neutral. In other words, valuing autonomy is a biased (non-neutral) value commitment. See for example, Allison Jaggar’s chapter on Liberal Feminism in *Feminist Politics and Human Nature*, pp.27-50.


5. See John Christman's Introduction in *The Inner Citadel*.

6. Dworkin argues that one must identify with one's desires. See: Dworkin "The Concept of Autonomy" reprinted in John Christman, ed., *The Inner Citadel: Essays on Individual Autonomy*, p.55. Frankfurt argues that one must decisively identify with one's desires for them to be authentic. See: Frankfurt "Freedom of the Will and the Concept of a Person" Ibid. p.71. There are problems with such accounts that cannot go unmentioned. First, there is a potential infinite regress of acts of identification. See John Christman's Introduction, Ibid., pp.9-11. Also, Dworkin changes his view. He no longer maintains that an individual must actually identify with his or her desires, due to the ambiguity of the notion identification; rather one must have the capacity to reflect and identify with one's ends.


8. Ibid., p.116.

9. Haworth, p.3

10. For Haworth, the attribution of value to one's ends is deeply connected with agency. See, Haworth: p.28, p.41-46, p.48, p.53, p.104.

11. Jeremy Waldron, *Liberal Rights* (New York: Cambridge University Press, 1993), p.162. I should note that although Waldron holds the view that to possess a conception of the good one must be autonomous, he doesn't think that this is in tension with a neutrality condition, because the neutrality condition does not require the state to be neutral among all possible conceptions of the good. I will return to this point in the third chapter.


15. I have in mind here a Benthamite utilitarian.


17. See endnote 9.

18. The extent to which a commitment to autonomy favors individual conceptions of the good over non-individualist conceptions is a deep and interesting question. This question has been the source of much of the liberal and communitarian debate. I will not, in this thesis, address this issue directly. My main concern is with the compatibility of a commitment to neutrality and autonomy, which I think is best answered by focusing on the concept of neutrality and what is demanded by liberals when we say the state must be neutral with respect to citizens’ conception of the good. I will address this issue in the third chapter.


Chapter 2

Rawls--The Political Move

A second strategy for constructing a liberal position immune to the kinds of criticisms articulated in the Introduction is to work out a merely political--and not comprehensive--conception of justice; this is the strategy developed by John Rawls and expressed most fully in his recent work *Political Liberalism.*¹ In order to understand both what is entailed in a political conception of justice and how such a conception purports to avoid any tension between the value of autonomy and neutrality, Rawls defines a political view negatively first. He does so in order to draw a contrast with non-political (or comprehensive views). A political conception of justice is not metaphysical on Rawls's view--that is, it does not assert its values beyond the political sphere. Rawls expresses this point in the following way:

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 free-standing view 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.²

As such, a political conception of justice is contrasted with what Rawls calls a "metaphysical" or "comprehensive" view in terms of the scope of the values within the conception of justice. A comprehensive moral doctrine has wide scope--that is, its content refers to and gives an account of what is valuable with respect to most or all of
human life.

Rawls characterizes comprehensive views in the following way:

[A moral conception] is comprehensive when it includes conceptions of what is of value in human life, and ideals of personal character, as well as ideals of friendship and of familial and associational relationships, and much else that is to inform our conduct, and in the limit to our life as a whole.³

Political liberalism, then, seeks to avoid any tension between the value of autonomy and a commitment to liberal neutrality by claiming that its commitment to the value of autonomy (or the conception of persons relied on) is limited to the political sphere.⁴ In other words, political liberalism claims not to be committed to the value of autonomy as a comprehensive moral value—a value in all spheres of life.

The success of political liberalism in providing an account of justice that is neutral among particular conceptions of the good life (i.e., comprehensive moral doctrines) is contingent upon the extent to which it accomplishes the following tasks: first, the scope of the doctrine must be limited to the political realm—that is, it specifies value only in the political sphere; and second, it must not appeal to the truth of any particular comprehensive moral doctrine to ground a commitment to its principles.

The extent to which Rawls succeeds in constructing a merely political doctrine, and thus in offering a cogent strategy for saving the doctrine of liberal neutrality, is an deep and important question. This question is the focus of my concern in this chapter. In an attempt to understand the limitations of Rawls's strategy I draw on the critique of political liberalism, put forward most notably by Jurgen Habermas, but more broadly, as
reflected by those working on a communicative approach to ethics.\textsuperscript{5}

I consider two challenges placed against the conception of persons upon which Rawls’s view relies. The first is that the conception of the person relies on the metaphysical claim that persons must in some sense be epistemologically and metaphysically prior to their ends; the criticism is that this at best violates neutrality and at worst is conceptually incoherent.\textsuperscript{6} The second challenge, if accurate, raises a serious and deep problem for political liberalism. At bottom, the claim is that Rawls builds normative concepts into the procedural structures (e.g., the original position), and in particular a non-neutral normative conception of persons, and this prevents the outcome--the principles of justice--from being neutral, from the point of view of acceptability among comprehensive doctrines. This will be explained in greater detail below.

I. Elements of Rawls's View

Rawls admits that the conception of a person relied on in his view is a substantive and normative one\textsuperscript{7}; this is permissible according to Rawls because it is a merely political conception. In other words, as long as substantive values are limited in scope they are permissible within a political conception of justice. Moreover, on Rawls’s view, the conception of the person relied on can be the focus of an overlapping consensus of reasonable comprehensive views, since it doesn’t require an appeal to the truth of any particular comprehensive doctrine. In order to both articulate and assess the criticisms levied against Rawls, it will be useful to first explicate a few central aspects of his view.
that purport to maintain political liberalism as a political and not comprehensive doctrine.

First, Rawls’s account of justice is a procedural one. In other words, there is no independent criterion for what counts as a just outcome. Rather, the justice or fairness of the outcome is defined in terms of the procedure from which it is derived. Although his account of justice is procedural, Rawls does not think it follows from this that the procedure itself must be neutral among citizens’ conceptions of the good life (comprehensive moral doctrines). Rather, he argues that political liberalism is neutral in its aim. A procedurally neutral doctrine would require a neutral justification of the political principles, so the very inclusion of any substantive normative conceptions in the procedure would transgress the requirements of neutrality. Such neutrality is not required for political liberalism, on Rawls’s view. Rawls argues all that is required is neutrality of aim, which “means that those institutions and policies are neutral in the sense that they can be endorsed by citizens generally as within the scope of a public political conception [of justice].”

Second, Rawls builds notions such as “the veil of ignorance” and “the original position” into the procedure as a means to ensuring the fairness of the outcome. The basic intuition behind these constructs is that we must limit the knowledge of the persons deliberating about just political principles in order to ensure a fair outcome. If people in the “original position” are behind a “veil of ignorance” (deprived of information about their particular identities) this ensures an impartial procedure and an impartial outcome. The principles of justice, so derived, will not privilege or promote any particular
conception of the good (comprehensive moral doctrine); they will be neutral among various conceptions of the good.

But, it may be argued, the use of the “original position” and “the veil of ignorance” appeals to a conception of persons and practical reason that is normative; and this introduces non-neutral substantive concepts into the procedure itself. Rawls thinks it is unproblematic to include substantive normative contents into the procedure insofar as these normative contents are merely political, which brings us to the third feature of Rawls’s view that needs further explication—namely, what is required for a conception of justice to be political rather than comprehensive.

There are three constituent features of a political conception of justice, according to Rawls. First, the focus of a political conception of justice is “the framework of basic institutions and the principles, standards and precepts that apply to it.” Second, a political conception of justice is a freestanding view. That is, it is modular; it is independent of any particular comprehensive doctrine, but an is essential constituent part of any comprehensive doctrine that is reasonable. Third, a political conception of justice as Rawls articulates it is a reflection of certain political ideals latent in western democratic culture, for example, the idea of society as a system of fair cooperation and the idea of citizens as free and equal persons.

The second aspect of a political conception relies on the idea of the possibility of an overlapping consensus. The notion of an “overlapping consensus” is analogous to a neutrality principle, and thus in a central component of the politically liberal view. Rawls
claims that an overlapping consensus “consists of all the reasonable opposing religious, philosophical, and moral doctrines likely to persist over generations and to gain a sizable body of adherence in a more or less just constitutional regime, a regime in which the criterion of justice is that political conception itself.” In other words, a conception of justice is political, then, if it can be the source of agreement among differing reasonable comprehensive moral doctrines. It is through the notion of an overlapping consensus that the commitment to the principles of justice is moral, rather than a mere prudential agreement among self-interested agents, on Rawls’s view.

One’s commitment to the principles of justice is moral because one accepts and endorses the principles of justice based on one’s own reasonable comprehensive doctrine, and one is committed to the values within one’s own comprehensive doctrine on moral grounds. Rawls contrasts a moral commitment to the principles of justice with an non-moral commitment based on self-interest--what Rawls calls a *modus vivendi*. A modus vivendi is a prudential agreement made by various self-interested agents who have conflicting conceptions of the good life, but lack the ability and sufficient resources to coerce, or otherwise force, opposing parties into abandoning their conflicting conceptions of the good (but would adopt that strategy--coercion--if it were not so costly). Given the interest each party has in not being frustrated by the others in seeking or promoting their ends, all parties seek some mutually advantageous agreement that sets limits on the actions of others.

Rawls must show that the political conception of justice, as he articulates it, can
be the focus of an overlapping consensus (and thus independent of the truth of any particular comprehensive doctrine), and yet still reflect a moral commitment. In order to assess the extent to which Rawls has articulated a political conception of justice, one need look closely at the notion of an “overlapping consensus” and the role it plays in the justification of such a conception. This is the focus of the second challenge.

II. Challenges to Rawls

I will consider the first challenge and Rawls’s response, before moving to the second challenge, because I think Rawls can respond to the former fairly easily while the second criticism poses a more serious challenge to his view. Nevertheless, the first criticism is one that needs be responded to. The main thrust of this criticism is that Rawls, through the model constructs of “the veil of ignorance” and “the original position,” places such strong epistemic constraints on the individuals deliberating about the principles that they cannot do the work which they were employed to do--namely, construct principles of justice. Under these epistemic constraints, persons have no means of individuating selves because they are abstracted from all particulars. Put differently, the argument is that without concrete information about the particular constitution of others’ identities--their histories, emotive-traits, etc.--one lacks sufficient information by which to pick out various persons--that is, to individuate selves. Without information about “other” particular selves, the deliberators lack the necessary epistemic tools to determine universalizable principles of justice, since the possibility of constructing
universalizable principles of justice, on Rawls’s view, requires that we imagine ourselves in the place of others.\textsuperscript{13}

This leads to conceptual incoherence, according to Seyla Benhabib, because the concept of a person relied on here requires that persons abstract from all particulars and this obviates the notion of individual “selves” altogether. The implication of Benhabib’s criticism is that, on Rawls’s account, persons must actually abstract from their ends. But this is mere interpolation since Rawls does not require such abstraction by persons constructing the principles of justice.

Rawls addresses this objection directly in \textit{Political Liberalism}. The criticism is misplaced insofar as it mistakenly attributes the requirement that persons actually abstract from their ends, as deliberators, to Rawls’s view. This interpretation fails to take into account that “the veil of ignorance” and “the original position” are mere constructs--devices of representation. Rawls writes:

\begin{quote}
I believe this [the claim that conception of a person requires selves to be metaphysically prior to their ends] to be an illusion by not seeing the original position as a device of representation. The veil of ignorance, to mention one prominent feature of that position, has no specific metaphysical implications concerning the nature of the self; it does not imply that the self is ontologically prior to the facts about persons that the parties are excluded from knowing. We can, as it were, enter this position at any time simply by reasoning for principles of justice in accordance with the enumerated restrictions on information.\textsuperscript{14}
\end{quote}

Surely, the latter claim is not something Benhabib would disagree with in spirit, for all it requires is a kind of reflective capacity on the part of persons--not complete abstraction from oneself. All Rawls requires is that persons at least be capable of
imagining themselves apart from some particular end at some particular time—not that persons could conceive of themselves independently of ends altogether. This is something that Benhabib is expressly committed to—namely, a “vision of the autonomous self” (See quote below). The question then becomes, does this vision of an autonomous self commit political liberals to relying on the truth of a comprehensive moral doctrine? This question is the focus of the second challenge.

The second criticism is the crux of the critique offered by communicative ethics. It is the claim that Rawls relies on a substantive normative conception of the person and that this requires an appeal to the truth of a comprehensive doctrine, which violates the neutrality principle. This criticism is expressed clearly and concisely by Seyla Benhabib.

Both liberalism and critical theory are committed to a vision of the autonomous self. Disagreements emerge around another issue: inasmuch as philosophical liberals try to reformulate conceptions of the self without involving assumptions drawn from a comprehensive theory but in the light of articulating publicly shared conceptions of personality and agency, they weaken their own moral vision and are led into conceptual incoherence.15

Benhabib concludes that through the attempt to rely on a “political conception of the person” Rawls both weakens his view and faces conceptual incoherence. We just considered the merits of the conclusion that through its conception of the person political liberalism is incoherent, and I concluded that this objection can be adequately responded to by Rawls, or defenders of his view. The second conclusion—that by attempting to avoid any claims to truth political liberalism weakens its own moral vision—is a more
forceful criticism, and the one I will consider in detail. In order to assess the validity of this criticism, we need look closely at the notion of the overlapping consensus to see if it can do the “work” which Rawls intends it to do--namely, serve to justify a commitment to political liberalism.

Habermas does just this; he considers the role that the notion of an overlapping consensus plays in the justification of political liberalism. In particular, Habermas examines “whether the overlapping consensus plays a cognitive or merely instrumental role: whether it primarily contributes to the further justification of the theory or whether it serves, in light of the prior justification, to explicate a necessary condition of social stability.”16 In other words, Habermas is interested in the role that the overlapping consensus is supposed to play in justification of Political Liberalism; he points to an ambiguity in the function of the overlapping consensus.

On the one hand, the overlapping consensus may be justified cognitively. Habermas, elsewhere17, defines a cognitive use of language: “In the cognitive use of language, by contrast, we thematize the content of the utterance as a proposition about something that is happening in the world (or that could be the case).” That is, “we raise a truth claim for the proposition asserted.”18 If the overlapping consensus plays a cognitive role in the justification of the theory, then it is either true or false that the overlapping consensus reflects a moral agreement as to the principles of justice. The purpose of the overlapping consensus may be to serve as a justification insofar as it reflects an actual consensus among citizens as to the constituent parts of a political doctrine; this is what I
interpret Habermas to mean by playing "a cognitive role." That is, the overlapping consensus may reflect an actual moral agreement about the principles of justice.

On the other hand, the overlapping consensus may play a purely instrumental role—that is, it may simply articulate a necessary condition of social stability. In Rawls's language, if the overlapping consensus plays an instrumental role in the justification of political liberalism, then it is a modus vivendi. That is, if the overlapping consensus reflects a modus vivendi, the it is a purely prudential agreement come to by self-interested agents who have an interest in not being frustrated by others and so agree to a mutually advantageous principles of cooperation. The crucial aspect of a modus vivendi, in this discussion, is that it is an agreement based merely on self-interest. It is not that the conflicting parties come to any agreement through the recognition of others rights or out of respect (i.e., moral reasons); rather it is simply the most prudent method for maximizing one's own self-interested goals.

Habermas points to this ambiguity in the role that the overlapping consensus may play in order to criticize Rawls's view. The structure of his argument is as follows: He articulates the disjunctive possibility—either cognitive or instrumental—for interpreting the role of the overlapping consensus as exhaustive of the possibilities. Next, he argues that Rawls's himself eliminates one half of the disjunct as an interpretation of the overlapping consensus—namely, the instrumental role—and thus this forces Rawls to claim that the overlapping consensus plays a cognitive role in justifying the political principles. And from this, Habermas argues that if the overlapping consensus is to play a
cognitive role, then the merely political nature as well as the neutrality of Political Liberalism is threatened insofar as it requires an appeal to the truth of its principles.

According to Habermas, Rawls rules out the possibility of interpreting the role of the overlapping consensus as merely instrumental one. To see this, consider the following: On a purely instrumental interpretation of the overlapping consensus the role of this notion is to articulate a necessary feature of social stability and stability is valued only out of self-interest. If the function of the overlapping consensus is a merely instrumental one, then the theory fails to be justified on any grounds other than pragmatic ones. But Rawls claims to have articulated a view that can be embraced on moral grounds by different comprehensive doctrines. That is, the commitment to the principles of justice is moral, in the appropriate way, only if the moral commitment is not justified by an appeal to the truth of any particular comprehensive moral doctrine. If this is not possible and the overlapping consensus merely reflects a necessary feature of a stable society--i.e., plays a merely instrumental role in the justification of the political conception of justice--then Political Liberalism fails and Rawls has shown only that the overlapping consensus is a modus vivendi. And we know from our description above that this is inadequate to justify a moral commitment to the principles of justice.

If we accept the disjunctive possibility that Habermas enumerates for the role that the overlapping consensus must play in justifying Rawls theory, then once the instrumental role is ruled out the only possible role left is a cognitive one. Habermas writes, "... if Rawls rules out a functionalist interpretation of justice as fairness, he must
allow some epistemic relation between the validity of his theory and the prospect of its neutrality toward competing worldviews being confirmed in public discourse.”

Habermas’s argument for this claim is as follows: The test of neutrality of normative concepts (e.g., the concept of a person) is not something that can be tested “immanently,” rather we must look to actual worldviews and determine if the normative concepts are indeed neutral. This is what he refers to as a test of acceptability. If the overlapping consensus is to play more than an instrumental (functional) role in the justification of political principles, then it must be possible that it be confirmed in actual public discourse. In other words, the test of acceptability is a test of the validity of the theory. Since the overlapping consensus reflects a moral commitment to the principles of justice, it plays a cognitive rather than instrumental role. If it plays a cognitive role, then it can be explained in cognitive terms, “that is, in terms of the confirmation of the assumption that justice as fairness is neutral toward ‘comprehensive doctrines.’”

The idea that the test of neutrality is a test of acceptability is clearly echoed in the pages of Political Liberalism. Rawls writes, “those institutions and policies are neutral in the sense that they can be endorsed by citizens generally as within the scope of a public political conception.” Habermas’s point is that this test of acceptability (neutrality) is something that must occur within actual public discourse, and can not be done in an immanent manner from within the theory. Once we move to actual public discourses in order to test the validity of the theory, Habermas thinks we are forced into making epistemic claims— that is, claims about the truth of the theory itself and the contents
within the overlapping consensus. Habermas’s criticism is just that Rawls cannot know that the conception of the person (or the other notions purported to be the focus of an overlapping consensus) relied on by Political Liberalism will be agreed to and reflect a moral commitment prior to consulting actual world views. He then thinks that once we consult actual world views as to what to include within an overlapping consensus we are forced into making claims about what in fact various world views are morally committed to; this then forces us into making truth claims about the contents agreed upon.

Rawls responds to Habermas’s claim that the overlapping consensus must play a cognitive justificatory role. His basic strategy is to disavow that the disjunctive possibility that Habermas claims the overlapping consensus must play is exhaustive. Rawls argues that there is a third kind of justificatory role that the overlapping consensus may play; that is, it is justified by an appeal only to political values in the public realm and is justified for particular citizens because it is grounded or embedded in their particular comprehensive doctrines. Since the political conception of justice can be embedded in various comprehensive doctrines it can be the focus of an overlapping consensus.

According to Rawls, political liberalism can avoid Habermas’s criticism because it can be justified pro tanto (with reference only to political values) and pro tanto justification does not require knowledge of actual comprehensive doctrines. Rawls claims that since political liberalism is justified pro tanto there is no need for a test of acceptability from the point of view of actual comprehensive doctrines. The kind of
justification Habermas attributes to political liberalism is a public justification and this is
different from the kind of justification required for an overlapping consensus. An
overlapping consensus requires that each individual can justify the political principles
from within his or her own (reasonable) comprehensive doctrine.\textsuperscript{22} This also explains
how the overlapping consensus can represent a freestanding political view, according to
Rawls. The values that are located in the political realm are embedded in various
comprehensive doctrines and not dependent upon the truth of any particular
comprehensive moral doctrine.\textsuperscript{23} Rawls thinks that justice should be conceived of as a
political and not metaphysical view. In other words, he wants to avoid claims of
universal truth. This follows from his view that "in a constitutional democracy the
public conception of justice should be, so far as possible, independent of controversial
philosophical and religious doctrines."\textsuperscript{24} In an attempt to avoid appealing to the truth of
controversial assumptions, Rawls introduces the notion "reasonable." The notion
reasonable substitutes for notions of truth, on Rawls's view. Rawls understands
"reasonable" as being predicable of persons and of comprehensive doctrines. There are
two criteria for reasonable persons. Persons are reasonable when they are willing to
accept both fair terms of social cooperation and when they are willing to accept the
burdens of judgement.\textsuperscript{25} The criteria for reasonable comprehensive views grows out of
the criteria of reasonable persons. Rawls identifies three features of any reasonable
comprehensive view. First, a reasonable doctrine is an exercise of theoretical reason.
Second, it is an exercise of practical reason. Third, a reasonable comprehensive view is
generally associated with, or draws upon, "a tradition of thought and doctrine."²⁶

In contrast to a comprehensive view, Rawls maintains that a merely political doctrine does not assert, or make claims about, the truth of its principles or the truth of its normative conceptions (e.g., the conception of a person). Rather, Rawls substitutes the notion "reasonable" for notions of truth. Political liberalism, then, asserts that the conception of persons, for example, relied on for the purposes of constructing principles of justice (or included within the overlapping consensus) is reasonable, rather than true, and is thus unproblematic.

However, as Habermas points out, Rawls does not provide an account of how "the reasonable" substitutes for "the true," he only tells us that it does. Rawls, then, owes us an account of the "reasonable" and the way in which it is supposed to substitute for the notion "true." He also needs to supply an account of how, as a result, this substitution absolves Rawls from asserting the truth of his doctrine (or the normative conceptions employed therein).

In the article in which Rawls directly responds to the criticisms made by Habermas, Rawls avoids expounding upon the notion "reasonable" and the role it is to play in the justification of political liberalism. In response he writes:

In answer to this (the question, how is the notion of the reasonable to substitute for the true), I have nothing to add beyond what has already been said. Political Liberalism does not use the concept of moral truth applied to its own political (always moral) judgements....I grant that the idea to the reasonable needs a more thorough examination than Political Liberalism offers.²⁷
The notion of the reasonable is the seminal concept upon which a political liberalism relies. If Rawls cannot articulate clearly and precisely the role it is to play in the justification of a merely political doctrine, this will pose a serious problem for the plausibility of his view. This is exactly my concern in the following section.

III. Evaluating Political Liberalism

I argue here that we have reasons for thinking that Rawls has not succeeded in constructing a merely political doctrine. First, though it may be that through the notion “reasonable” Rawls avoids appealing to the truth of any particular comprehensive doctrine, the notion at least serves the same functional role (or purpose) as the concept (or predicate) “true.” As a result, it is not clear that political liberalism avoids relying on controversial assumptions in the appropriate way. To see this, consider the following: The notion “reasonable” allows that particular comprehensive moral doctrines are excluded from the scope of political liberalism. These doctrines are excluded in virtue of the fact that they do not possess the features (or meet the criteria) of “reasonable” views.

Now it is not clear where the notion reasonable comes from, on Rawls’s view. But, it is the case, for the purposes of political liberalism, the criteria for reasonable persons and reasonable comprehensive moral doctrines are fixed. In other words, the criteria represent unchanging standards by which to judge a particular view as falling within the scope of political liberalism or being excluded from the scope. Moreover, persons or views are reasonable in virtue of the fact that they share certain assumptions.
This does the same work as asserting the truth of these assumptions insofar as it allows us to exclude particular comprehensive moral doctrines in virtue of the fact that they do not share a commitment to the values located in the "political realm."

Rawls avoids claiming that the conception of the person, and other "merely political conceptions," are true because he wants to avoid the relying on the truth of controversial assumptions as a basis for his conception of justice. It is not clear, though, that the use of the notion "the reasonable" obviates this difficulty. If this is coupled with the fact that Rawls's avoids offering a procedurally (or justificatory) neutral account of the normative contents within political liberalism, then political liberalism begins to look like a comprehensive doctrine.

This can be seen through Rawls's reliance on the idea of a shared notion of practical reason. He writes:

We adopt, then, a constructivist view to specify the fair terms of social cooperation as given by the principles of justice agreed to by the representatives of free and equal citizens when fairly situated. The bases of this view lie in the fundamental ideas of the public political culture as well as in citizens' shared principles and conceptions of practical reason.29

One question that comes immediately to mind is, if we (citizens) have shared conceptions of practical reason, why do we disagree about the correct comprehensive moral doctrine? After all, is it not the case that our conception of practical reason determines our conception of morality? Put simply, practical reason is the faculty by which we decide what we ought to do. The force of the rhetorical question just posed is this: Insofar as Rawls claims that citizens possess shared principles and conceptions of practical reason it
would seem that if we can agree about the principles that govern what we ought to do, then for any given dilemma we would agree about what we ought to do. This, of course, is assuming that we can agree about which principle is relevant to a given situation, identify the relevant features to pick out the appropriate principle, etc. But given that we have “shared conceptions of practical reason,” it seems like we would have no problem doing so.

There are two implications of this criticism. The first is that if citizens have shared conceptions of practical reason, and if that notion were filled out (i.e., given any content), then that content would be subject to scrutiny—that is, we ought to be able to assess the validity of that content. So at the very least it looks like Rawls has to assert the truth of the conception of practical reason his view relies on. Moreover, it seems likely that any such notion of practical reason will be controversial. Second, given that there needs to be content to the notion of practical reason employed, I argue, that once the content is supplied the conception itself will tell us how to address moral issues. In other words, I see no reason in principle why our “shared practical reason” would limit itself to the political sphere.

Rawls attempts to avoid this implication by claiming that the conception of the person and, by implication, the conception of practical reason in political liberalism is a merely political conception. What is required for such a conception to be merely political according to Rawls is that:

[P]ersons can accept this conception of themsevles as citizens and use it when
discussing questions of political justice without being committed in other parts of their life to comprehensive moral ideals often associated with liberalism, for example, the ideals of autonomy and individuality. 30

In other words, Rawls thinks there is space to include certain substantive values and conceptions within a limited sphere (the political realm) without thereby committing oneself to any particular comprehensive doctrine. This is just what it means to construct a political conception of justice on Rawls's view. What Rawls needs to do in order to respond to the sorts of concerns I am raising here is provide a principled account of ways in which we limit the scope of the values needed for a political conception of justice to the political sphere. In other words, he needs to draw a sharp line between the political realm, and what is located therein, and what counts as a comprehensive doctrine. He attempts to do just this by employing the notion "reasonable."

The content of the political realm is distinguished from a comprehensive moral view, according to Rawls, because we do not assert the truth of the conception of a person, for example, that political liberalism relies on. Rather, Rawls merely claims that such a conception is "reasonable." There is, perhaps then, space for Rawls to respond to the sorts of concerns Habermas and I raise, and that is to fill out the notion "the reasonable" in further detail and give an account of exactly the ways in which it is distinguished from the notion true. But, this, as Rawls admits, is further work that needs to be done.

The goal of the preceding is not to present decisive reasons for rejecting either Rawls's view or the strategy of constructing a merely political liberal doctrine. Rather,
the goal is to point both to places where there are tensions within the view as well as places where the view needs to be filled out in more detail. Given the difficulties within political liberalism, I would like to suggest that the project of constructing a comprehensive liberal doctrine ought not to be abandoned. What we need to do at this point, then, is offer an account of how a liberal committed to autonomy as a comprehensive value can maintain a this commitment consistent with the demands of a neutrality condition.

To contrast such comprehensive views with Rawls’s project consider two obvious instances of comprehensive liberal doctrines as found in the political philosophies of Kant and Mill. Both philosophers are committed to the value of autonomy and their political principles grow out of this commitment. More contemporary examples of comprehensive liberal doctrines can be found in the works of Jeremy Waldron and Ronald Dworkin. At bottom, both Waldron and Dworkin argue that liberalism requires a commitment to the value of autonomy. From this it follows that, as liberals, Waldron and Dworkin are committed to a conception of the person as an autonomous agent. By Rawls’s standards, both of these views rely on the truth of controversial assumptions.

This is unpromblematic for liberalism, on Waldron’s view. He argues that although liberals are committed to a principle of neutrality among various conceptions of the good, the justification of this commitment does not itself have to be neutral. Moreover, Waldron argues that the view that liberals must offer a neutral justification of liberalism leads to incoherence. I will address these issues, in depth, in Chapter Three.
Whether or not Rawls agrees with Waldron’s conclusion that the claim that liberals ought to be neutral with respect to the justification of a neutrality principle “is simply a recipe for incoherence,” he at least shares the view that liberals are not required to offer a procedurally neutral account. This is why he thinks that it is not a problem that political liberalism relies on a normative conception of the person. Where he differs substantially from Waldron, or other comprehensive liberals, is in his view that reliance on these normative conceptions is permissible only if they are both not required to be true and can be the focus of an overlapping consensus. In other words, these normative contents must be capable of being affirmed by “reasonable” comprehensive views and the focus of a “reasonable overlapping consensus.”

As we have seen, the notion “reasonable” appears to do the same work as the concept “true.” And so, it is not clear that Rawls has avoided relying on the truth of controversial assumptions in the way required for a merely political doctrine. At the very least then, Rawls owes us a further account of the notion “the reasonable” and the role it is to play in the justification of political liberalism. In the absence of this account, I would like to turn and consider how liberals committed to the value of autonomy as a comprehensive value—that is, as a value in all spheres of life—might maintain this commitment in conjunction with a neutrality condition.
Notes to Chapter Two:


2. Ibid. p.10.

3. Ibid. p.13.

4. Seyla Benhabib uses the phrase “autonomous self” to refer to the conception of persons that liberals employ. I take these notions to be synonymous and will generally use the latter for the purposes of this chapter since it is the language frequently used by Rawls. The reader should note that one could substitute “autonomous self” for “concept of the person” at any point in the discussion that follows.


6. Ibid. p.166.

7. Political Liberalism, p.18, fn.20.

8. Ibid. p.192.


11. Ibid. p.12 and p.144.

12. Ibid. p.15.

13. Seyla Benhabib writes, “I conclude that a definition of the self that is restricted to the standpoint of the generalized other becomes incoherent and cannot individuate among selves. Without assuming the standpoint of concrete other, no coherent universalizablity test can be carried out, for we lack the necessary epistemic information to judge my moral situation to be “like” or “unlike” yours. Benhabib, “The Generalized and Concrete Other,” in Women and Moral Theory, ed. Eva Feder Kittay and Diana T. Meyers (U.S.A.: Rowman & Littlefield, 1987). pp. 163-167.


19. Ibid.


21. Rawls, “Reply to Habermas,” in *The Journal of Philosophy*, Vol. XCI, No.3, March 1995, p.142. Rawls defines pro tanto justification in the following way: “That is, the political values specified by it can be suitably ordered, or balanced, so that those values alone give a reasonable answer by public reason to all, or nearly all, questions concerning constitutional essentials and basic justice. This is the meaning of pro tanto justification.”

22. Ibid. p.143.

23. Ibid.


27. Rawls, “Reply to Habermas” p.150.

28. Though, he does say that he thinks it originates in Kant’s work. “The distinction between the reasonable and the rational goes back, I believe, to Kant: it is expressed in his distinction between the categorical and the hypothetical imperative in the *Foundations* and his other writings. The first represents pure practical reason, the second represents empirical practical reason.” *Political Liberalism*, p.48-49 fn. 1.

29. Ibid. p.97.


Chapter 3

Neutrality and Autonomy Reconsidered

Up to this point I have considered two strategies for avoiding any tension between autonomy and neutrality. I have attempted to point to some weakness of both such strategies. At this point, I want to take a closer look at the concept of neutrality to in order to understand exactly what a commitment to a neutrality condition requires of a liberal political philosophy. I suggest a way in which, as liberals, we can consistently maintain a commitment to a neutrality condition in light of a commitment to the value of autonomy. We merely need to qualify what is required by such a condition.

I. The Concept of Neutrality

I begin this discussion, then, with a distinction that prevails in the literature devoted to the issue of neutrality in liberalism. The neutrality condition is interpreted, generally, in one of two ways—that is, as requiring neutrality of outcome or neutrality of justification. Though there is a lack of consistency with respect to the labels used to refer to this distinction, the general intuition behind the distinction remains the same. For instance, Will Kymlicka refers to the distinction as one between "consequential" and "justificatory" neutrality.¹ According to Kymlicka, justificatory neutrality demands that "the state does not justify its actions by reference to some public ranking of the intrinsic value of different ways of life, for there is no public ranking to refer to."² In contrast,
consequential neutrality "requires that the state seek to help or hinder different life-plans to an equal degree--that is, government action should have neutral consequences."³

Kymlicka cites Joseph Raz’s *The Morality of Freedom* as the source for this distinction between consequential neutrality and justificatory neutrality. However, those phrases are absent from Raz’s discussion. Raz refers to the distinction as one between "neutral political concern" (i.e., what Kymlicka calls consequential neutrality) and "exclusion of ideals" (i.e., Kymlicka’s justificatory neutrality).⁴ John Rawls makes a similar distinction, though his distinction is between “procedural neutrality” and “neutrality of effect.”⁵ For Rawls, procedural neutrality is similar to what Kymlicka calls neutrality of justification. He writes:

One way [to define neutrality] is procedural, for example, by reference to a procedure that can be legitimated, or justified, without appealing to any moral values at all. Or if this seems impossible, since showing something justified appears to involve an appeal to some values, a neutral procedure may be said to be one justified by an appeal to neutral values, that is, to values such as impartiality, consistency in application of general principles to all reasonably related cases...⁶

In contrast to this way of characterizing neutrality, Rawls points out that one can think neutrality of effect or influence is demanded by liberalism; this is analogous to Kymlicka’s consequential neutrality. Neutrality of this sort requires the outcome of some political principle or action to be neutral among various conceptions of the good life.

Rawls rejects the idea that liberalism demands neutrality of effect, as do the majority of liberals. I will consider the arguments for this claim below.

In any case, though the terms may differ the intuition behind the distinction
between justificatory neutrality and consequential neutrality (these are the terms I will use throughout my discussion) remains the same. There is a sense in which one might think that the role of the state is to ensure a neutral outcome—that is, where no particular conception of the good is privileged or promoted as a result of some political principle or political action. In contrast to this characterization of neutrality, one might think that in order to remain neutral among citizens’ particular conceptions of the good life, the political principles which govern the state must be justified without appealing to the value of any particular conception of the good. This characterization of justificatory neutrality is somewhat vague and misleading. I intend, through the following discussion, to gain a clearer understanding of what is required of liberalism when we say that liberals must provide a justificatory neutral account of their political principles.

With that caveat, I am concerned here with the sort of liberalism that holds the view that the neutrality condition requires justificatory neutrality. Interpreting the neutrality condition as requiring consequential neutrality is unappealing to most liberals for several reasons. Jeremy Waldron, for example, argues that defining the neutrality condition as requiring neutrality of consequences faces the theoretical difficulty of hypothesizing some baseline standard by which to judge or measure the diverse effects of some state action. Therefore, one would need to identify some common standard across differing conceptions of the good life by which the consequences of various state actions ought to be measured. The possibility of hypothesizing such a standard requires both that differing conceptions of the good not be incommensurable and that we can identify some
feature by which they are commensurable. The possibility of identifying such a feature, theoretically, which is uncontroversial seems quite difficult. Moreover, Waldron argues, that in addition to the theoretical difficulties of postulating such a standard, demanding consequential neutrality would face the practical difficulty of requiring legislators to predict the effects of any given policy on various conceptions of the good life.

Will Kymlicka argues that “two fundamental components of liberal justice--respect for liberty and fairness in distribution of material resources--both preclude consequential neutrality.” In support of this claim, Kymlicka refers to the familiar arguments offered in discussions of egalitarian distributive justice, namely, those arguments concerning expensive tastes. Briefly, the argument is that the state ought not to compensate those with voluntarily adopted expensive tastes by providing them with additional resources, because doing so would violate the demands of fairness. The standard example mentioned to illustrate the unfairness of compensating those who voluntarily adopt expensive tastes is as follows: Imagine two people, one who prefers beer and the other prefers champagne. Imagine further that the person who prefers champagne cultivated this taste, voluntarily, to appear “sophisticated.” If the state were required to ensure that the distribution of resources achieved neutral consequences among these two people, then the person who prefers champagne would have to be given more resources (because champagne is more expensive than beer). This example is supposed to illustrate the intuition that such unequal distribution is unfair, because the person who chooses champagne could choose to prefer beer and the beer-lover is treated deferentially
as a result of choosing less expensive tastes.

An important assumption relied on in this argument is that citizens are held responsible for their conceptions of the good; so it is assumed that citizens are capable of reflecting upon and revising their conceptions of the good in light of the resources they can expect to be fairly provided to them. If resources are then distributed equally, despite varying tastes, then it might turn out that someone with less expensive tastes than others would be able to use their resources to acquire more goods. This is unproblematic, according to the argument just described, so long as the distribution was fair (i.e., equal) in the first place. This argument applies to the issue of neutrality because the conclusion is that despite non-neutral outcomes, the state ought to distribute resources equally among citizens. Moreover, on this basis, attempting to create neutral outcomes would violate the conditions of a fair distribution.

In addition, if one thinks neutrality of consequences is required, one would allow for substantial state intervention to ensure the appropriate “neutral outcome.” So, for example, if the market were left to operate and the consequences of the market were non-neutral, the state might interfere with a particular citizen’s life (and the way he or she chooses to live) in order to produce a neutral outcome, and this may involve substantial violations of autonomy. It seems likely, then, that a defender of such a view would likely run into conflict between this kind of neutrality and a commitment to the value of autonomy. A commitment to the value of autonomy demands a certain sort of respect for that value—or a kind of non-interference. If consequential neutrality were demanded
of the state, then the practice of enforcing this kind of neutrality might conflict with
valuing autonomy insofar as persons' autonomous choices compromise a neutral
outcome. So for the purpose of exploring the possibility of reconciling the purported
tension between neutrality and autonomy, I will focus my attention to those accounts of
liberalism that demand justificatory neutrality.

II. Neutral Justification of the Neutrality Condition

The notion of justificatory neutrality has been interpreted by some as the
requirement that liberal must offer a neutral justification of the neutrality condition.
Charles Larmore, for example, argues that “if liberals are to follow fully the spirit of
liberalism, they must also devise a neutral justification of political neutrality.”\textsuperscript{10} I want to
consider this claim in detail, for I take it to be equivalent to the claim that liberals must
remain neutral among all possible conceptions of the good\textsuperscript{11}, and if Larmore is right, it
looks like the possibilities for reconciling any tension between a commitment to the value
of autonomy and the neutrality condition are grim. We know this based on the
arguments offered in chapter one, namely, that a commitment to the value of autonomy,
as a value that the state ought to promote in its citizens, is non-neutral with respect to
non-autonomous conceptions of the good life. In fact, Larmore argues that because of
this tension between autonomy and neutrality, we ought, as liberals, abandon a ·
commitment to the value of autonomy as a political ideal. Appeals to controversial
values such as autonomy are inadequate to ground a commitment to neutrality, thinks
Larmore, and this, in part, is the reason that liberals must offer a neutral justification of the neutrality condition.

Larmore argues that traditional liberal justifications of the neutrality condition of liberal principles, such as those found in the work of Mill and Kant, which appeal to some value such as autonomy or experimentalism in order to ground a commitment to neutrality themselves violate the conditions of neutrality. Therefore, he thinks that in order to avoid appealing to any controversial value (i.e., one that is itself non-neutral) liberals must offer a neutral justification of neutrality. He goes on to offer an account of how this might be done. I focus on that account for two reasons. First, I intend to assess the extent to which the project of offering a neutral justification of the neutrality condition is successful. Second, I am interested in the accuracy of his criticisms of the project of grounding a commitment to neutrality in autonomy.

Larmore argues that in order to be neutral about being neutral, a liberal view must not appeal to any controversial claims in order to justify its political ideals, commitments, or policies. The model he offers to ensure this sort of neutrality is based on a “universal norm of rational dialogue.” The basic intuition behind this is that in the face of disagreement participants in conversation ought to “retreat to neutral ground” in order to justify their claims. In other words, participants in some particular discourse must, in order to offer a neutral justification, abstract from the disputed claims in the conversation and retreat to neutral ground. Larmore readily concedes that this is not the only possible avenue, that is, we are not driven to this view by the demands of, say, rationality. For
example, one might react to such disputes by removing oneself from the conversation. So in order to convince us that we ought to continue the conversation and retreat to neutral ground, Larmore does two things. First, he appeals to persons’ sense of community and desire for civil peace as a reason for continuing the conversation. Secondly, he provides an argument from the idea that all persons are entitled to equal respect.14

Larmore argues that “the neutral reason for continuing the conversation with them [other citizens] must lie in the wish to show everyone equal respect.”15 Larmore claims that persons are due equal respect in virtue of their capacity to possess a conception of the good life. However, he claims that this does not reduce to valuing their capacity for autonomy.16 In this discussion, Larmore realizes that he is invoking substantive normative claims when he appeals to the values of community, civil peace, or showing persons equal respect, and in doing so, he “retreats” a bit on his earlier claim that one must offer a neutral justification of neutrality.

Therefore, he weakens the criteria for what a neutral justification of neutrality amounts to, and claims that his account is “largely neutral.”17 Larmore realizes that he is excluding particular conceptions of the good, namely, those that do not, in particular, value community, or don’t accept claims to equal respect; but now he claims that justificatory neutrality does not require the state to be neutral among all possible conceptions of the good. He writes, “It does not require that the state be neutral with respect to all conceptions of the good life, but only with respect to those actually disputed in society. Where everyone agrees about some element of human flourishing, the liberal
should have no reason to deny it a role in shaping political principles." Larmore’s argument now has quite a different color than before, however. By the end of his discussion, it does not appear that Larmore has maintained a commitment to the view that a liberal must provide a neutral justification of neutrality as he characterized it earlier.

I explore this shift in Larmore’s argumentation and suggest that the reason he ends up providing a not so neutral justification of neutrality is because it is not possible to do so. If the claim that liberals must provide a neutral justification of neutrality implies that the justification lacks any substantive normative commitment, then appealing to neutrality, as such, is insufficient to ground a commitment to the neutrality condition. Beyond this, I suggest further that the reason Larmore, and others like him who think that justificatory neutrality entails being neutral about the neutrality condition, makes this mistake because of an ambiguity with respect to what the notion of “justification” refers.

In his discussion of the requirements of neutrality, Larmore rules out three traditional arguments for liberal neutrality. They are: skepticism, appeals to the value of experimentation, and appeals to the value of autonomy. These sorts of arguments are ruled out, according to Larmore, because they are not arguments all reasonable persons can accept. In other words, they are not neutral among various conceptions of the good life, for example, because not everyone values autonomy or not everyone thinks the proper response to skepticism is neutrality. It is this argument that convinces Larmore that a neutral justification of neutrality is demanded by the liberal tradition. He claims that it is possible to reject those controversial justifications just mentioned and still offer a
justification of the neutrality condition that regards neutrality as the fundamental political
value.\textsuperscript{25} However, Larmore does not succeed in offering such a justification. Rather, he
simply introduces new reasons that one ought to be committed to the neutrality condition.
He appeals to the values of community, cooperation, and respecting persons to ultimately
justify a commitment to a neutrality condition.

So we can see that it is not that we value neutrality simpliciter, for its own sake.
Rather, it is in light of our other value commitments that we think neutrality is a good
thing. We must appeal to some substantive normative value in order to justify a
commitment to the neutrality condition. The claim that one must offer a neutral
justification of the neutrality condition is either incoherent (because it requires that we
justify a normative claim without providing a substantive reason) or an appeal to the
value of neutrality. In either case it is not the kind of justification Larmore thinks he has
provided.

Jeremy Waldron argues that the idea that the neutrality condition demands a
neutral justification is incoherent.\textsuperscript{26} Waldron argues that because the neutrality condition
is a normative claim that is purported to be action-guiding, it requires justification other
than an appeal to neutrality itself. In other words, we must offer some reason, other than
an appeal to neutrality, for thinking the state ought to remain neutral among citizens' conceptions of the good life. Waldron offers a brief analysis of the concept of neutrality;
this analysis is limited, according to Waldron, because the concept of neutrality can be
fully articulated only in light of its justification.
According to Waldron, analysis of the concept of neutrality reveals the following:

First, the concept of neutrality presupposes a contest between two or more sides. In other words, the concept presupposes pluralism, of some sort. Second, Waldron claims that the concept of neutrality implies the existence of some third party, a party not directly involved in the dispute at hand, that is required to be neutral. It is interesting to note that this is not the way Larmore conceives of the notion, for he thinks that the participants in a particular dispute can be required to be neutral. Thus, his inclusion of the notion of "retreating to neutral ground" is an example of the requirement that participants in a conversation may need to be neutral.

The third feature of the concept of neutrality revealed by analysis, according to Waldron, is that there are a range of possible actions available to those that are required to be neutral. The possibilities are located somewhere between the third party taking part in the dispute on the same terms as those involved and the third party being completely uninvolved. Waldron offers these two criteria as fixed points of reference for the filling out the concept of neutrality, since each of these possibilities represents one extreme or another of non-neutrality. Beyond these three conditions, the concept of neutrality is a contextual notion that is only filled out further when one's particular reasons for desiring neutrality are explained.

If we follow Waldron's analysis of the logical requirements of the concept of neutrality, then it is clear that nowhere in that analysis are we required to be neutral about the neutrality condition. This being the case, then, we can look at the claim that we must
be neutral about the neutrality condition as one possible justification among many. Put differently, justifying a commitment to the neutrality condition by appealing to neutrality is not structurally different than appealing to, say, the value of autonomy. To be clear, let’s consider the adequacy of such a justification.

If one claims that the proper justification of the neutrality condition is neutrality itself, then the structure of that claim is as follows: The role of the state is to remain neutral among citizens’ conceptions of the good life, and this is so because neutrality is itself a value, or in Larmore’s words is “the fundamental political value.” But we may ask whether or not it is true that neutrality is a fundamental political value, whether the reason a liberal demands that the state remain neutral among various conceptions of the good life is that neutrality is a fundamental political value?

One can respond by arguing that it doesn’t make sense to appeal to neutrality as the reason for remaining neutral among citizens’ conceptions of the good life. Waldron put this point a bit stronger and argues that to require a neutral justification of neutrality is incoherent. This is exactly right, if one interprets the claim that to be neutral about the neutrality condition one must in some sense be indifferent or impartial with respect to the neutrality condition. For as we said above the neutrality condition is a substantive normative claim, the force of which one cannot be impartial or neutral about. Though if one interprets the claim that to be neutral about the neutrality condition rests on an appeal to the value of neutrality, as it appears Larmore does, then one is making a substantive claim about the value of neutrality.
So it looks like Larmore makes an appeal to at least three different values in order to justify a commitment to neutrality—namely, a sense of community, equal respect, and neutrality itself. We can recall that Larmore's reasons for excluding appeals to the value of autonomy to ground a commitment to the neutrality condition were that the value is controversial or non-neutral itself. I think, though, it is clear that the grounds which Larmore himself appeals to do not meet this requirement either. It is not the case that everyone agrees to some uncontroversial sense of community or that everyone thinks neutrality is a value in itself.

In light of this analysis, it is clear that one must invoke some value, or reason, for grounding a commitment to the neutrality condition. It is still an open question, however, whether one can justify a commitment to a neutrality condition based on the value of autonomy. Before considering this question--of whether the neutrality condition can rest the value of autonomy--I make a distinction to keep in mind while considering the merits of different justifications of a neutrality condition.

There are at least two ways in which one might consider the concept of neutrality, and in order to be clear, we ought keep them separate. First, one might consider what the neutrality condition demands of a liberal state. And second, one might consider the reasons that justify a commitment to the neutrality condition. That is, one might ask why ought we be committed to a neutrality condition; why is neutrality a good thing?

It may turn out that in the process of the analysis of the concept of neutrality, we may consider the reasons we think the neutrality condition is required of us (the state), but
this is to ask a different question. It is to ask on the basis of what do we justify a
commitment to the neutrality condition. I suggest that the failure to be cognizant of this
distinction in discussions of the requirements of the neutrality condition results in the
confused idea that liberals must be neutral about the neutrality condition.

What I suggest, then, is two things. First, the language used frequently in
discussions of liberal neutrality, for example the phrase “justificatory neutrality,” is in
part responsible for the idea that the neutrality condition demands a neutral justification.
In other words, the failure to recognize the distinction just mentioned is a result of
conflating the analysis of the concept of neutrality with its justification. The second thing
I argue, then, is that the answer to the question about the basis for neutrality cannot be
neutrality itself. In light of this argument, I suggest, in the final section of this chapter, a
way of cashing out the notion of justificatory neutrality that is at least consistent with
valuing autonomy, or appealing to the value of autonomy to ground a commitment to
neutrality.

III. Autonomy and Neutrality

I argue, in this section, that our commitment to the neutrality condition makes best
sense in light of a commitment to the value of autonomy. In other words, it is because we
value persons’ capacity for autonomy--their ability to govern themselves--that we think
the state is required to be neutral among various conceptions of the good life. Moreover,
it is in light of our commitment to autonomy that we set the scope or parameters for the
concept of neutrality. To make this clear consider the fact that it is exactly in those cases in which persons' autonomy (or capacity for autonomy) is violated that we think the state is required to interfere, or to \textit{not} remain neutral.

For example, imagine a case in which someone is coerced, say, into signing a contract and wants to challenge the legitimacy of that contract. Surely, the state does not stand by neutrally or impartiality when such coercion occurs. We do not think the proper response from the state to a plea to remove oneself from such a coerced contract is to claim that its job is to remain neutral among citizens' conceptions of the good life, even when one citizen's conception (coercing others) infringes upon another's conception of the good (to live uncoerced). Rather, there are avenues, set up by the state, for such victims to remove themselves from contracts to which they do not authentically consent. I suggest that it is the lack of authenticity or autonomy in such cases that compels the state to interfere, on the one side rather than the other. In other words, the state would, in such a case, support the victim of coercion rather than the coercer, and I suggest this support is a result of valuing autonomy.

Or we might consider mandatory educational laws, for example: the state is not neutral among those conceptions of the good that prohibit, say, the education of women while the state is neutral with respect to the particular means by which one might educate one's children. Parents may send their children to public schools, private schools, or home-school them. The state is non-neutral in these cases because we realize that a certain level of education is required for autonomy, for example, education promotes the
sort of critical thinking and reflection needed to be autonomous. Though there is no set
formula that the state enforces as to the manner of the education, so long as such formulae
meet certain broadly defined criteria that do not defeat autonomy, the state is neutral with
respect to the manner in which one is educated. It is cases such as these that illustrate the
dependency of the neutrality condition on the value of autonomy.

What these examples illustrate is that our commitment to the value of neutrality is
conditioned upon our commitment to the value of autonomy. Where autonomy is at
stake, as in the case of education, I suggest that we no longer think that neutrality among
all possible conceptions of the good is demanded by the state. More strongly, I suggest
that it would be wrong for the state to stand by neutrally and allow persons’ autonomy or
capacity for autonomy to be neglected or violated.

I am not arguing that the only way to justify a commitment to the neutrality
condition is by appealing to the value of autonomy, though I am arguing that it is at least
consistent and plausible to do so. Moreover, I am arguing that in light of the fact that we
must appeal to some value to ground a commitment to the neutrality condition, it makes
best sense of our liberal intuitions, if the value we appeal to is that of the autonomy of
citizens. The above examples are meant to illustrate such intuitions. Having said this, I
would now like to reconsider Larmore’s claim to ground a commitment to a neutrality
condition in the demand for equal respect, in order to illuminate my argument about the
relation of autonomy and neutrality.

Larmore suggests that the “neutral reason” that liberals are committed to the
neutrality condition lies “in the wish to show everyone equal respect.”24 This respect for persons comes from our recognizing every persons’ capacity for “coherently developing beliefs from within their own perspective.”25 However, Larmore claims such respect for persons does not reduce to valuing autonomy. He adds the caveat that “from within their own perspective” is “not necessarily on their own, or autonomously.”26

I disagree with Larmore’s claim here, for I think the best way to make sense of our commitment to the neutrality condition is to appeal to the value of autonomy—thatis, the value of persons’ capacity of coherently developing beliefs from within their own perspective. What is valuable about this capacity is that beliefs which result from careful reflection which represent one’s own perspective are genuine and authentic, in other words they are autonomously held. If we think in terms of the concept of autonomy as articulated in the first chapter, I argue that it is in light of one’s relationship to one’s ends that we can distinguish some ends as more valuable than others, or at least this relationship provides us with a reason to be neutral with respect to some ends and not others.

This is not, however, to retreat to Haworth’s view, as articulated in the first chapter. Recall that Haworth argues that one has a conception of the good only if one’s ends, goals, or projects are autonomously held. I deny that claim, as I argued in response to Haworth, for this rules out a lot of persons and what count for them a conceptions of the good. Also, Haworth attempted to ground a commitment to the neutrality condition in an appeal to the general subjectivity of value, and as we saw this is not consistent with
valuing autonomy as he characterizes it. However, it is still open to me to argue that, although one’s non-autonomous ends do in some sense count as conceptions of the good, one’s autonomous ends are preferable, or more worthy, than one’s non-autonomous ends in virtue of the fact that they are autonomously held. And moreover, the recognition of the value of autonomy, as such, makes best sense of a commitment to a neutrality condition.

In order to make this claim, I must offer some principled reason for thinking that neutrality is constrained by the value of autonomy. The best I can do is to point to cases in which persons lack autonomy, and argue that there is something of fundamental value missing in such cases. And it is exactly in those cases where is this value is lacking that we question the role of neutrality. The value of autonomy is found most explicitly in those cases where we demand the end of oppression and domination; autonomy is a fundamental value in any society that purports to be free and democratic, that is, any liberal society.

The oppression of women is a good test case, for example, to illustrate the value of autonomy. Imagine a case in which a particular woman is subordinate, say, to her husband (I say imagine, but we certainly do not have to imagine such cases, there are pervasive examples in contemporary culture). For our purposes, let’s imagine that this particular woman is financially dependent upon her husband’s income, and exemplifies the stereotypical housewife; in other words, she does not work outside the home, is responsible for maintenance of the house and care of the children. Moreover, let’s
imagine that this woman does not reflect on her position, but assumes, as she was taught, that the role of a woman is to perform such duties. It is not enough that she simply does not actually reflect on her situation to defeat her autonomy. It must be that as a result of conditioning her ability to reflect is hampered and that it does not even occur to her to reflect on her status as a housewife.

Now I certainly do not mean to suggest with this example that any particular woman could not authentically desire to be a “housewife.” I simply mean to suggest in those cases where there is a lack of reflection and authentic acceptance of one’s position, something of value is missing. When women, and men for that matter, call for the end of such oppression of women they are appealing to the value of autonomy. What is missing in the case of the housewife just mentioned is that she lacks, in some sense, the ability to reflect on her status as a housewife and the ability to hold that desire as a genuine.28

It certainly is not the case that everyone will agree with me when I say autonomy is a fundamental value, or that what is missing of value in the cases of oppression and domination is lack of autonomy. For example, Iris Marion Young argues against oppression and domination, but does not do so on the basis of the value of autonomy. She defines oppression and domination in the following way:

Oppression consists in systematic institutional processes which prevent some people from learning and using satisfying and expansive skills in socially recognized setting, or institutionalized social processes which inhibit people’s ability to play and communicate with others or to express their feelings and perspective on social life in contexts where others can listen. While the social
conditions of oppression often include material deprivation or maldistribution, they also involve issues beyond distribution. Domination consists in institutional conditions which inhibit or prevent people from participating in determining their actions or the conditions of their actions.29

Young argues that oppression and domination hinder emancipation and justice, though emancipation is not defined in terms of individual autonomy as we've considered the notion. However, I would like to suggest that oppression and domination, as Young describes them above, are negative phenomena because they impede autonomy.

Consider domination: domination, on Young's account, prevents people from determining their actions or the conditions of their actions. This can clearly be read in terms of the value of autonomy. Autonomy is valuable, in part, because persons who are autonomous govern their own lives. Autonomy is defined as self-determination. Accordingly, when such domination occurs, one's ability to determine oneself—to be self-governing—is the value at stake.

Oppression is a little more difficult to illustrate, but I think it can still be understood in terms of the value of autonomy. When one calls for the end of certain sorts of oppression one is opposed to a kind of systematic interference with persons and with those capacities which allow them to interact within particular kinds of social settings.

Young points to the deprivation of certain sorts of material goods, perhaps food will serve as an example, as resources required to develop those capacities relevant to developing skills, communicating effectively, etc. I would argue that what is important about these resources is that they are necessary for the development of one's cognitive capacities
(e.g., critical reflection) that in turn are necessary for autonomy.

When persons are oppressed they lack the ability to determine their lives as they would otherwise, absent the oppression, choose to do so. It is the focus on the absence of those capacities that permit an individual to live his or her life in such a way as to develop his or her own perspective and the ability to communicate that perspective to others, which seem to me to underscore the value of autonomy.

Moreover, I argue that in cases where such oppression and domination occurs, the state ought not remain neutral due its recognition and respect for the value of autonomy. The examples of the coerced contract, childrens’ education, and the subservient housewife serve to illustrate this point. I am clearly not being neutral with respect to the value of autonomy, and I argue that the state ought not to be either. Moreover, I argue that our commitment to the neutrality condition makes best sense in light of an appeal to the value of autonomy, for it is just in those cases where autonomy is violated that we think our commitment to the neutrality condition is constrained.

There are two features of this account worth highlighting for our later discussion of perfectionism (Chapter Four). First, the notion of autonomy employed here is similar to the concept presented in Section I of Chapter One. On that account autonomy is defined in terms of the relation one has toward one's ends; it is a threshold concept. One is either autonomous with respect to some end or one is not, one cannot become more or less autonomous with respect to any particular end. Second, due to understanding autonomy as a threshold concept the state similarly is required to respect autonomy when
it is present, and respect is a threshold concept as well. That is, the state either respects the value of autonomy or it does not for any particular case it cannot provide more or less respect. These features will be important in order to contrast the sort of liberalism I am defending here with a perfectionist view and thus are worth mentioning now.

IV. Justificatory Neutrality and Autonomy

The question that we now face is whether an appeal to the value of autonomy to ground our commitment to the neutrality condition is consistent with the demands of justificatory neutrality. I argue that it is. To see this consider that we now know that justificatory neutrality does not demand a neutral justification of the neutrality condition, for this is incoherent. Rather, we must supply some reason (i.e., some justification) for our commitment to the claim that the role of the state is to remain neutral among citizens’ conceptions of the good life. I have attempted to argue that one such justification is to appeal to the value of autonomy.

I argue further that in the absence of any value or reason supporting a commitment to neutrality, we are at a loss to explain what is required by a neutrality condition. It is only in light of the value that we appeal to in grounding a commitment to the neutrality condition (e.g., autonomy) that we are able to articulate the scope or parameters of that principle. In light of this justification, the neutrality condition, then, requires that the state be neutral with respect to citizens’ autonomous conceptions of the good. In other words, neutrality, of the sort considered here, demands that the state not
discriminate among citizens’ particular conceptions of the good life based on the content of those conceptions; rather, the state may discriminate based on the manner in which it is held (i.e., is it autonomous?).

It does not follow from this claim that the state ought to go about systematically interfering with non-autonomous conceptions of the good. For there may be other considerations which rule out such behavior. For example, it may be quite costly to do so, or it might substantially hinder the potential autonomy of citizens if the state were to act in such a way. So while conceiving of the neutrality condition in this way does not necessarily allow for state intervention with respect to non-autonomous conceptions of the good, it does allow for state promotion of autonomy.

The view that I have articulated here is consistent with the analysis of the concept of neutrality borrowed from Jeremy Waldron. Waldron’s analysis of neutrality revealed the following criteria for its justification: (1) We must provide some reason for thinking neutrality is desirable, and we cannot be neutral about the force of this reason; (2) The concept of neutrality presupposes pluralism—that is, a contest between two or more sides; (3) The concept of neutrality implies that there is some third party, a party not directly involved in the dispute at hand, that is required to be neutral; and (4) There are a range of possible actions available to those that are required to be neutral, located somewhere between the third party taking part in the dispute on the same terms as those involved and the third party being completely uninvolved.

Grounding a commitment to the neutrality condition in autonomy certainly meets
the first condition--namely, a reason is provided for remaining neutral among citizens' autonomous conceptions of the good life. It is just because we think that it is valuable for citizens' to live their own lives, to be self-governing, that we think the state ought not to interfere with the autonomy of its citizens. It is exactly in those cases where one's autonomy is threatened that we think the state is compelled to interfere or abandon its neutral stance. The second and third criteria are met at least in the following respect: I argued, in the first chapter, that there are a plurality of values on this account; there is a plurality of conceptions of the good life, some autonomously held and some not, and these may or may not conflict. The fourth condition simply lays out a range of possibilities for filling out the requirements of the neutrality condition further, the extremities of which include the third party (the state on my account) remaining uninvolved, and the third party involving itself at the same level of the participants. I think that the account presented here meets this condition insofar as it avoids both of these extremes. It is clear that on the account just presented the state must not remain uninvolved when the value of autonomy is at stake, and the state doesn't involve itself on the same level as participants insofar as it is seen as an arbitrator or an authority to which one might appeal.

I think we are now in a position to reflect on those definitions of justificatory neutrality as presented in the first section of this chapter. Kymlicka defined justificatory neutrality as demanding that "the state does not justify its actions by reference to some public ranking of the intrinsic value of different ways of life, for there is no public
ranking to refer to." John Rawls recognizes the difficulty of demanding a completely neutral justification of the neutrality condition, and on this basis claims the one must appeal to neutral values. As quoted earlier, he writes, "If this [requiring a neutral procedure that appeals to no moral values] seems impossible, since showing something justified appears to involve an appeal to some values, that is, to values such as impartiality..." What I have attempted to argue throughout this chapter is that impartiality (or neutrality) is itself a non-neutral value, at least in the sense that it is not neutral among all possible conceptions of the good. Moreover, to argue that the state must be neutral about the neutrality condition is incoherent.

The notion of justificatory neutrality, as articulated by those discussed in the first section of this paper, is meant to contrast with consequential neutrality. In that section we considered liberal arguments against consequential or outcome neutrality, and I think they are sound ones. However, I think, Kymlicka and Rawls's way of cashing out the notion of justificatory neutrality sounds much like demanding a neutral justification of the neutrality condition. But as we have seen, this is incoherent. So in light of that consideration I would like to consider what we can mean, as liberals, by the notion justificatory neutrality. To put this differently, if, as liberals, we are looking to articulate some conception of neutrality that does not reduce to neutrality of consequences, what is that conception?

We have come to several important conclusions in the process of this discussion the question just mentioned. Justificatory neutrality cannot mean liberals must be neutral,
where this is equivalent to involving no substantive normative claims in that justification.

In other words, liberals cannot be required to be neutral among all possible conceptions of the good. Our commitment to the neutrality condition must be informed by some substantive normative commitment. I suggested that the value of autonomy is one possible justification of a commitment to the neutrality condition. I suggested, further, that it is in light of such a substantive justification that we set the scope or parameters of what it is that we (the liberal state) is required to be neutral about. So in this sense, I think, the best way to characterize this kind of neutrality is as a “principled neutrality.” In other words, we must offer some principled reason of why liberal state ought to be committed to a neutrality condition. But such principled neutrality cannot be filled out in absence of substantive normative commitments. Yet, this kind of neutrality stands in contrast to a neutrality of outcome or consequences.

The worry over the tension between a commitment to the value of autonomy and the neutrality condition arises when we conceive of the neutrality condition as requiring the state to remain neutral among all possible conceptions of the good. But as we have seen this cannot be what is required of liberalism, because being neutral about the neutrality condition simply is not possible. Once we understand this, we are in a position to realize that we must invoke some substantive conception of the good in order to justify a commitment to the neutrality condition. I argue that autonomy is the value to which we ought, as liberals, appeal.

Now the question is whether this makes liberalism a perfectionist doctrine, that is,
a doctrine which articulates some ideal of human flourishing and thinks it is the states job to promote that conception. In other words: Does the fact that the neutrality condition is grounded in terms of the value of autonomy make the view just articulated perfectionist in a problematic sense? In the final chapter, I address this question.
Notes to Chapter Three:


2. Ibid. p.886.

3. Ibid. p.883.


6. Ibid. p.191.


9. Raz as a special case.

10. Larmore, p.53.

11. Although I should note that Larmore denies this equivalence. All I mean, here is that being neutral among all possible conceptions of the good is equivalent to offering a neutral justification insofar as neither can invoke any controversial substantive normative claim in order to justify a commitment to the neutrality condition. I will focus on this issue later in my discussion.

12. Larmore, p.55.

13. Larmore, p.54.


15. Larmore, p.61.

16. See Larmore, p.63. I will return to this point later in my discussion, for I think this does reduce to valuing persons’ capacity for autonomy.
17. Larmore, p.61.


19. For example, see Bruce Ackerman's *Social Justice in the Liberal State* (New Haven, CT: Yale University Press, 1980).


25. Larmore, p.63.

26. Larmore, p.64.

27. This is not to say that the state ought to go about systematically uprooting particular non-autonomous housewives from their homes. For the cost of such action (monetary or personal) may outweigh the concern for autonomy in any particular case. I would suggest the state has a responsibility to mitigate, perhaps through education and awareness programs, that conditioning which hinders autonomy.

28. I say her status as a housewife, because I take it that autonomy is a local rather than global concept. In other words, when we speak of persons’ autonomy we must do so with respect to particular ends, and not the person as a whole.


30. See endnote 3, above.

31. See endnote 5, above.
Chapter 4
Perfectionism?

Liberalism is traditionally contrasted with perfectionism. A perfectionist doctrine is one that articulates some ideal of human flourishing, and thinks it is the state's job to promote that ideal in its citizens. Aristotle’s emphasis on the life of contemplation as the "good human life" is a classic example of such a perfectionist doctrine. A perfectionist state, in contrast to the kind of liberalism just articulated, could allow for substantial state intrusion into citizens’ lives if such intrusion were oriented toward realizing the ideal of a flourishing life.

Liberalism, as a political philosophy, is located within a tradition which is, at the least, unsympathetic to perfectionist views. In other words, liberalism is defined by a tradition that says the state ought to leave it to citizens to decide what counts as a good life. Perhaps the most eloquent expression of this anti-perfectionist sentiment is found in John Stuart Mill’s *On Liberty*, where he enunciates the harm principle. Mill writes:

> That principle is, that the sole end for which mankind are warranted, individually or collectively, in interfering with the liberty or action of any of their number, is self-protection. That the only purpose for which power can be rightfully exercised over any number of a civilized community, against his will, is to prevent harm to others.¹

Gerald Dworkin claims that an appeal to the value of autonomy is often used as a reason for opposing any perfectionist political morality. He argues:

> As a political ideal, autonomy is used as a basis to argue against the design and

81
functioning of political institutions that attempt to impose a set of ends, values, and attitudes upon the citizens of a society. This imposition might be based on a theological view, or secular visions of a good society, or the importance of achieving human excellence along some dimension of human achievement...In particular, then, autonomy is used to oppose perfectionist or paternalistic views.²

In addition, one can easily read anti-perfectionism into liberal calls for the protection of individual freedom or toleration, or for that matter into the neutrality condition itself. The appeal of a neutrality condition is that it limits the extent to which the state may legitimately interfere with citizens' lives. Though as we have seen, this limitation is not absolute. As I argued in Chapter Three, the neutrality condition is constrained by the value of autonomy. That is, the neutrality condition is justified in terms of the value of autonomy.

Some critics of liberalism, as well as some liberals, worry that through its commitment to the value of autonomy an autonomy-based liberalism becomes a perfectionist doctrine insofar as the value of autonomy may rest on an ideal of human flourishing. Isaiah Berlin, for example, in his seminal essay, "Two Concepts of Liberty," argues in favor of a negative conception of liberty over a positive one (e.g., autonomy) because, historically, tyranny and oppression have been justified as a means to greater emancipation. Berlin puts the point in the following way:

Once I take this view [defining freedom positively], I am in a position to ignore the actual wishes of men or societies, to bully, oppress, torture them in the name, and on behalf, of their 'real' selves, in the secure knowledge that whatever is the true goal of man (happiness, performance of duty, wisdom, a just society, self-fulfillment) must be identical with his freedom—the free choice of his 'true', albeit often submerged and inarticulate, self.³
The worry over perfectionism is two-fold. First, the worry is that through a commitment to the value of autonomy, the sort of liberalism considered here commits itself to some ideal of human flourishing (a conception of the best human life), which is problematic insofar as liberals think this is just the sort of thing the state ought to remain silent about. Second, the worry is that if a commitment to the value of autonomy commits the liberal state to some controversial ideal of human flourishing, then promotion of autonomy is the promotion of that controversial ideal of the good human life.

It is essential to the project of articulating a coherent view of autonomy-based liberalism that such worries be mitigated. My goal, then, in this final chapter is to reflect on the view of liberalism as articulated in the third chapter and to contrast it with a perfectionist doctrine. In order to do so, I set out two tasks. First, I consider a perfectionist justification of a principle of non-interference in order to contrast it with a liberal justification of a principle of non-interference on behalf of the state (a neutrality condition). In order to draw out this contrast I mention, briefly, the criteria of a perfectionist morality and consider the implications of extrapolating that morality to the political realm. Second, I argue that it is precisely because certain central features of a perfectionist morality are absent from the sort of liberalism being defended here that this autonomy-based liberalism fails to be a perfectionist political doctrine.
I. Perfectionist Morality Applied to Political Institutions

Thomas Hurka attempts to articulate a philosophically defensible account of perfectionism as an account of morality. Although Hurka provides a detailed account of what human perfection amounts to, I will not concern myself with the details or plausibility of his view as an account of morality. Rather, I draw on Hurka as a resource for identifying the more general and abstract criteria of perfectionist views.

A perfectionist view has four main components: First, any perfectionist view must rest on some ideal of human nature. Second, based on this view of human nature, we can articulate some ideal of human good that is based on assigning intrinsic value to those essential properties which make up human nature. The third feature of a perfectionist view is the claim that developing those essential properties which make up human nature constitutes the good human life. Based on this conception of the good, Hurka’s account of morality is such that moral (or good) actions are those which maximize the good (i.e., perfection). Accordingly, then, Hurka’s perfectionism is a consequentialist theory. In other words, good actions are those which produce the best outcomes (i.e., maximize perfection). The fourth feature of a perfectionist view is that it is agent-neutral. That is, perfectionism tells “us to care not just about our own perfection but about that of all humans.”

Based on his perfectionist view of morality, Hurka attempts to articulate a view of political arrangements or institutions. On Hurka’s view, the relationship between perfectionism as a moral theory and perfectionism as political theory is that we just to
extend the moral claims into the political realm. He characterizes the relationship as follows:

Given its [a perfectionist morality] agent neutral maximizing structure, it supports the following standard for political evaluation: The best political act, institution, or government is that which most promotes the perfection of all humans. Moreover, Hurka argues that this way of conceiving the role of political institutions is compatible with liberalism; Hurka describes his view as both perfectionist and liberal egalitarian. However, I think his view is not appropriately liberal.

Hurka appeals to the value of autonomy as an intrinsic value, and moreover, as a perfectionist ideal. He claims that based on the intrinsic value of autonomy a perfectionist is in a position to argue in favor of non-interference on behalf of the state. In other words, Hurka suggests that one can make an argument for liberty from seeing autonomy as a perfectionist ideal. Insofar as he provides an argument for protection of liberty, Hurka thinks he provides a liberal view. Implicit in this claim is the idea that liberalism is defined in terms of a commitment to liberty. Though this is in some sense accurate, there is more to a liberal view than a mere commitment to the value of liberty--defined here as autonomy. Below, I argue that beyond a commitment to the value of liberty (or autonomy) liberal views are distinguished from non-liberal view based on the justification of that commitment. Before I enter upon this argument, I consider Hurka’s justification of non-state-interference in its citizens’ lives.

Hurka argues that a commitment to some principle of non-interference by the state can be justified from within a perfectionist view that sees the realization of autonomy as a
perfectionist ideal. The justification of this principle of non-interference is couched in terms of the high cost of systematic interference. Systematic interference by the state in its citizens' lives is unlikely to achieve the goal of promotion of autonomy, since to be autonomous one must in some sense bring about the condition of self-governance oneself. To put this point differently, autonomy as a perfectionist ideal is best cultivated in an environment in which the state is committed to a principle of non-interference. Hurka puts this point in the following way:

Government interference with self-regarding action reduces citizens' autonomy and especially their deliberated autonomy. At the same time, it rarely succeeds in promoting other perfections and can work in several ways to diminish them, by removing routes to excellence, inducing less valuable motives, and weakening self-direction.\textsuperscript{10}

Although Hurka endorses non-interference by the state as a means to promote autonomy, he contrasts this with a commitment to a neutrality condition, as thought to be demanded by some liberals. He thinks that the neutrality condition demands that the state remain neutral with respect to all conceptions of the good. He writes:

Although perfectionism supports liberty, it does not go beyond it to support the stronger ideal of state neutrality. According to this ideal, the state must not only not coerce citizens to make them better, it must never aim, coercively or otherwise, to promote one set of values over others. It must be neutral about the good, never having as its justification for acting that some ways of life are intrinsically preferable to others.\textsuperscript{11}

He goes on to argue that such a condition is not demanded by the liberal tradition. I agree with this conclusion insofar as it refers to the neutrality condition as he has characterized it in the passage quoted above. He defines the neutrality condition as the requirement that
the state abstain from promoting any values—that is, the state must remain neutral with respect to any value or good. We considered just this sort of claim in Chapter Three. In a sense, Hurka as well as those liberals we considered in Chapter Three, who think that liberals must provide a neutral justification of the neutrality condition, are relying on an implicit dichotomy between a neutral justification of the neutrality condition, on the one hand, and perfectionism, on the other hand. For example, Charles Larmore thinks, in part, the reason that a neutral justification of the neutrality condition is demanded by the liberal tradition is because he thinks that to justify the neutrality condition in terms of some value is to be committed to promoting a conception of the good life, which he thinks liberals must not do.12

The assumption relied upon here that is that in the process of picking out some value by which to ground a commitment to the neutrality condition liberalism becomes perfectionist. However, it is not the case that in the absence of providing a neutral justification of the neutrality condition liberalism becomes a perfectionist doctrine. This dichotomy is a false one. There is space between being neutral “all the way down” and offering a perfectionist doctrine, and it has to do with the sort of justification offered for a commitment to the neutrality condition.

II. Liberal Justification of the Neutrality Condition

As I argued in the Third Chapter, the neutrality condition does not require a
neutral justification—that is, we are not required as liberals to be agnostic about what
gives value to particular conceptions of the good. It is incoherent to say that one must
provide a neutral justification of the neutrality condition, for the neutrality condition is a
substantive normative commitment, about which one cannot be neutral. Rather, it is in
light of our commitment to some value that we think neutrality is demanded of the liberal
state, and I suggest autonomy as that value. Although the justification of the neutrality
condition is in terms of the value of autonomy the autonomy-based liberalism I am
presenting does not thereby become a perfectionist view. Once we consider the ways in
which Hurka's view is illiberal, we will be in a position to consider the ways in which the
view of liberalism suggested in Chapter Three is anti-perfectionist.

A perfectionist state that is committed to the value of autonomy thinks it the
state's job to promote the value whenever it can. Hurka argues that a principle of non-
interference can be justified from within such a perfectionist view. The justification for
non-interference rests on the claim that non-interference is the best means to promotion
of the end--namely, autonomy. So a perfectionist state asserts the value of autonomy and
is constrained in promoting that value only insofar as the promotion is self-defeating. In
principle, then, there is no reason why the perfectionist state committed to the ideal of an
autonomous person as the ideal human life has to constrain itself in the promotion of that
ideal, so long as the promotion achieves its goal. In other words, a perfectionist state will
impose its ideal of human flourishing up to the point of diminishing effects. Non-
interference is justified in terms of pragmatic concerns.
There is, I think, space to contrast the sort of liberalism articulated in Chapter Three with this kind of perfectionism. Although the autonomy-based liberalism defended here includes a commitment to a neutrality condition that rests upon a commitment to the value of autonomy, it does not thereby reduce to a perfectionist view. One important contrast is the sort of justification given for non-interference (or neutrality). The sphere of acceptable interference on behalf of the state in its citizens' lives is constrained, on a liberal view, not because interference would inhibit a goal of promoting autonomy, but because respecting autonomy as a value requires neutrality among autonomous conceptions of the good. It is not the states', but individuals', business to determine what counts as the good human life. The liberal state is concerned only that those choices which reflect citizens' conceptions of the good life are authentic (i.e., autonomous) preferences—so that citizens are indeed self-governing. As I pointed out in Chapter Three, the sort of liberalism I am defending here is concerned with the manner in which citizens' hold their conceptions of the good and not the content of those conceptions.

What distinguishes a "liberal" view from a "perfectionist" view, then, is the kind of justification offered for a commitment to a neutrality condition or a principle of non-interference. The perfectionist offers a pragmatic concern of diminishing effects, where in principle there is no limit to permissible state interference so long as the ideal human is promoted as a result. In contrast, I suggest that a liberal justification is distinguished insofar as acceptable interference on behalf of the state in its citizens' lives is restricted due to respecting, or recognizing, the value of autonomy.
Since the justification of the neutrality condition rests upon the value of autonomy, the limit of legitimate state interference is couched in terms of the demands of respecting autonomy. Even if it turned out that the state could successfully extend its powers to develop “flourishing” citizens, on a liberal account it would be wrong to do so—not based on pragmatic concerns, but moral ones. Though the sort of liberalism defended here is absent of such an account of human flourishing. As I argued in Chapter Three, on the conception of autonomy we are concerned with, what is important is the manner in which a particular end is held—not the content of that end.

As identified above, a perfectionist morality entails some view of human nature as well as the claim that there are essential properties constituted by that nature that ought to be promoted. Moreover, the promotion of these properties is constitutive of the “good human life.” Human nature on a perfectionist account is teleological. Furthermore, based on its description of human nature a perfectionist view makes the normative claim that to fulfill one’s nature is to be acting or functioning well—that is, to flourish. The sort of autonomy-based liberalism considered here lacks any such ideal of human nature, where that ideal reflects a normative commitment to any particular conception of the good life.

The absence of an account of human nature prevents liberals from identifying any essential properties from which to define “the good life.” The kind of liberalism presented in the Third Chapter remains silent with respect to the worth of differing autonomous conceptions of the good; it lacks a standard by which to rank them.
Autonomy-based liberalism, as I have articulated it here, is concerned only with the manner in which citizens hold their conceptions of the good; it is indifferent to the content of those conceptions.

The account of autonomy presented in Chapter One is offered as a description of the features necessary to count as "self-governing." Beyond that descriptive account, I consider what normative claims about the value of autonomy one might make. These normative claims, however, were couched in terms of the relation one might have toward his or her ends and not the content of those ends. To count as autonomous, then, one need merely stand in the appropriate relation to one's ends--that is, have the capacity for critical reflection, acquired in the appropriate way, and hold them authentically.

Autonomy is defined in terms of a process, a psychological attitude, or a relation; it is not an end state. As Gerald Dworkin puts it, autonomy can be seen as "a formal notion (where what one decides for oneself can have any particular content)."¹³ John Christman further elucidates the notion of a "content-neutral conception of autonomy." He writes:

> One implication of this [his] theory of autonomy is that people could turn out to be autonomous despite having desires for subservient, demeaning, or even evil things and life styles. I don't take this to be a defect of the model. It only reveals that the concept of autonomy we are discussing is 'content-neutral.'¹⁴

On this account of autonomy, then, there is no particular conception of the good by which to identify the good human life.

On such a content-neutral account of autonomy the notion is seen as a threshold
concept. One is either autonomous with respect to some end—that is, has the capacity to reflect on that end, acquired it in the appropriate way, and holds it genuinely—or one is not autonomous with respect to some end. One cannot become more or less autonomous with respect to any particular end. As I said above autonomy is not seen as an end state. Accordingly, seeing autonomy as such has implications for the way in which we think autonomy as a value ought to be promoted, which in turn has implications for state action.

I suggest that on such a content-neutral account of autonomy the state is required to respect or recognize the value of autonomy, but that this is different than the kind of promotion a perfectionist state would allow. A perfectionist account of autonomy would define that notion in terms of some essential properties of human nature, rationality for example. On such an account the promotion of autonomy might include selecting certain autonomously held ends over others in virtue of the fact that they are more rational or contribute to the overall promotion of autonomy. In other words, one can been seen as more or less autonomous in virtue of the ends that one hold, and this may have implications for state action insofar as the state sees its job as that of promoting autonomy full stop.

In contrast, the liberal state, as I am defending it here, does not understand autonomy in this way and as a result differs in the ways in which it thinks it appropriate to esteem that value. Autonomy merits a kind of respect, and respect, as such, is a threshold concept so that one (or the state) either respects another or does not; there is no more or
less respect. Accordingly, when the liberal state promotes autonomy it respects that value, and beyond offering that respect it is neutral. Once respect is given there is no more promoting to be done. This contrast with a perfectionist state where in principle there is no limitation to the possible promotion of autonomy.

In addition, I argued in Chapter One that although autonomy is a value, and in some sense confers value on some ends, we ought not define conceptions of the good in terms of the value of autonomy. To do so would rule out the possibility of the satisfaction of preferences as legitimate conceptions of the good as well as, perhaps, some relational conceptions of the good. Seeing these sorts of ends, goals, or projects as conceptions of the good requires the acknowledgment of other values besides autonomy. Moreover, in acknowledging the value of some such conceptions of the good one abstains from defining the ideal or value of autonomy as the sole end to which a good human life ought to be directed. Though I do argue in Chapter Three that one ought to see autonomy as a fundamental value, the liberalism I am defending does not thereby become a perfectionist morality.

III. Conclusions

Throughout this thesis I have attempted to articulate a coherent view of autonomy-based liberalism--that is, a view which is committed both to the value of autonomy and to a neutrality condition and yet not a perfectionist view. My suggestion is that liberalism, as such, is best understood as seeing its commitment to the neutrality condition as
constrained by the value of autonomy. However, understanding the neutrality condition in terms of the value of autonomy does not reduce to a perfectionist view, because the justification of that condition is done in term of respecting the value of autonomy (conceived of as a threshold, or content-neutral, concept). This reconciliation of the commitment to neutrality and autonomy offers the best interpretation of liberalism’s theoretical commitments and it does so consistently.
Notes to Chapter Four:


5. Ibid. This summary is also draw from a paper by Thomas Christiano, “Problems of Perfectionism,” presented at the Central Division Meetings of the American Philosophical Association.


Bibliography


