

THE UNIVERSITY OF MANITOBA

AUTONOMY, THE AUTHORITY OF REASON, AND REASONABLE AUTHORITY:
KANTIAN ELEMENTS IN CONTEMPORARY POLITICAL THEORY

by

DENNIS KLIMCHUK

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE
OF MASTER OF ARTS

DEPARTMENT OF PHILOSOPHY

WINNIPEG, MANITOBA

August, 1990



National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-71766-1

Canada

*AUTONOMY, THE AUTHORITY OF REASON, AND REASONABLE AUTHORITY:
KANTIAN CLEMENTS IN CONTEMPORARY POLITICAL THEORY*

BY

DENNIS KLIMCHUK

A thesis submitted to the Faculty of Graduate Studies
of the University of Manitoba in partial fulfillment of the
requirements of the degree of

MASTER OF ARTS

(c) 1990

Permission has been granted to the LIBRARY OF THE UNIVERSITY
OF MANITOBA to lend or sell copies of this thesis, to the
NATIONAL LIBRARY OF CANADA to microfilm this thesis and to
lend or sell copies of the film, and UNIVERSITY MICROFILMS to
publish an abstract of this thesis.

The author reserves other publication rights, and neither the
thesis nor extensive extracts from it may be printed or
otherwise reproduced without the author's written permission.

"Reason has no dictatorial authority; its verdict is always simply the agreement of free citizens, of whom each one must be permitted to express, without let or hindrance, his objections or even his veto."

Kant, Critique of Pure Reason

"Covenants, without the Sword, are but Words."

Hobbes, Leviathan

Table of Contents

Introduction	3
Chapter One: The Kantian Picture	10
Chapter Two: Kantian Autonomy Naturalized	49
Chapter Three: Autonomy and Procedure	85
Chapter Four: Autonomy and Personhood.....	153
Conclusion	190
Appendix: Kant Defended Against Stroud and Rorty	203
Bibliography	215

INTRODUCTION

Contemporary political theory appears to be in a state of disarray. There seems to be little agreement about what political theory ought to do, and what it can hope to achieve. Traditional divisions amongst political theorists along ideological lines have deepened or been replaced by divisions at the meta-level. Foucaultians, Critical Theorists, Libertarians, Rawlsians, and so on, are divided not so much (or not only) over traditional questions about the structure of the just polity, but over questions of the nature (and possibility) of the grounding of political critique.¹ The reception of Rawls's Theory of Justice illustrates this well. Originally heralded as a return to traditional normative political theory, it was met in the literature most often by criticism not of the substantive theory of justice it contained, but by criticisms of the elaborate justificatory scheme Rawls had developed to ground his theory of justice.

Inasmuch as this thesis is concerned with political meta-theory, Rawls's book, and its reception, set the climate for the

¹The distinction I am drawing is a matter of degree. I am not suggesting that meta-level concerns have been introduced into political theory only recently, but rather that such concerns have been given an atypical amount of attention within the last few years. Concerns that would have, until recently, be relegated to the philosophy of natural and social science have inserted themselves into political theory. The philosophy of science itself seems to be experiencing a similar preoccupation with its own methodology, as evidenced by the increasing attention paid to methodological naturalism.

concerns that I address here. Perhaps more significantly, in this context, the publication of A Theory of Justice served also to introduce Kant as an important figure in contemporary political theory. Since its publication (though likely not entirely due to its publication), scholarship in Kantian political theory has grown significantly. However, I do not plan to review much of this scholarship, nor to assess Kant's explicit political theory. Rather, I wish to employ Kant largely for his utility in helping us to sort out the bewildering variety of approaches to political theory currently available in the philosophical marketplace. In this thesis, I will argue that within the apparent disunity in contemporary political theory, a certain unity can be found along both historical and theoretical lines. Specifically, I will argue that voices within contemporary political theory as apparently disparate as John Rawls, Jürgen Habermas, Richard Rorty, Karl-Otto Apel, and Michel Foucault, can be understood as responding to certain questions left to us by Kant.

This claim, in itself, is relatively non-controversial. All the philosophers I have mentioned have, to a greater or lesser degree, articulated their positions in terms of their relationship to Kant. This is most explicit in Rorty and Apel, who often use their self-diagnosed relationships to Kant as pithy statements of their positions on the philosophical map. Much of my argument, however, is more a matter of reconstruction than catalogue. Rather than taking the self-diagnosis of these

theorists to be exhaustive statements of their relationship to Kant, I take these moments of reflection on their positions in the Kantian tradition as clues to a more comprehensive reading of contemporary political theory. As such, I am concerned more often with the assumptions and implications of these contemporary political theorists than with their stated critical and positive agendas.

Kant is not a transparent figure, and the theorists with whom I am concerned often disagree not only in terms of their relationship to Kant, but on the interpretation of Kant against which they articulate their own positions. Central to the defence of my thesis is an interpretation of Kant that is in itself controversial. I resist certain popular trends in the literature concerning the question of what it is we ought to read in Kant's corpus in an attempt to articulate his political theory. Of his explicit writings in political theory, I treat only one, a brief paper entitled "What is Enlightenment?" My interpretation of Kant's politics--as evidenced by my reliance on this essay, and recourse to certain passages in the first two Critiques--emphasizes the general Kantian question of what authority reason may be said to have, and how this authority is to be grounded. I argue, in Chapter One, that this interpretation suggests that we not read Kant so much as a philosopher who had, among his offerings, a political theory, but rather that we recognize a political element to the entire critical philosophy. The Kantian picture that emerges from this

reading--and, as I argue in the remainder of the thesis, still has a hold on contemporary political theory--suggests that there is something of an interface between epistemology and politics: that reason may be said to have authority if certain epistemic conditions (impartiality) are secured, *and* if certain political conditions (freedom) are secured. In other terms: 'the authority of reason' is secured only when both certain *de jure* and certain *de facto* conditions obtain. At the center of this Kantian picture are two key claims: (1) The methodological question of the grounding of radical political critique and the substantive question of the nature of autonomy are inexorably linked, and (2) Autonomy has both a transcendental and a political element, neither of which can be reduced to the other. This second claim essentially recasts the *de jure/de facto* link introduced above: in Kant's terms, autonomy is both noumenal and phenomenal, that is, a question both of free will and the political conditions that must obtain for the free will to be exercised in the political realm.

But the Kantian picture contains what A.C. Genova calls a 'philosophical liability': Kant's transcendental idealism and the reliance of this position on the noumenal/phenomenal distinction. As I will argue in Chapter Two, both Rawls and Habermas have at the core of their respective meta-ethics an attempt to naturalize Kantian autonomy so as to save the possibility of ethics from the need to posit the noumenal realm. The basic move in this naturalization is a re-interpretation of what Kant called the

'pathological' motivation of the will. Rawls reinterprets the pathological substantively, identifying it with the intrusion of morally arbitrary knowledge in the process of moral deliberation. Habermas reinterprets the pathological procedurally, identifying it with ideological forces that distort the procedure of consensus-formation. However, neither Rawls nor Habermas can complete this naturalization without 'transcendental residue.' That is, as Habermas admits explicitly, and Rawls implicitly, the neo-Kantian account of autonomy must take recourse to the transcendental. So at a basic level, the neo-Kantian account of autonomy remains faithful to Kant, insofar as autonomy is still understood by neo-Kantians to have both political and transcendental elements.

In the third and fourth chapters I consider criticisms that approach the neo-Kantian account of autonomy first from the transcendental and then from the political side, respectively. In a nutshell, I argue that the former approach fails, and that the latter succeeds. I develop my defence of the neo-Kantian recourse to transcendentalism regressively, that is, I begin with a strong position--Kant's original formulation of transcendental arguments--and pare it down to a weaker position--Habermas's argument for the transcendental preconditions of communication. So my argument approaches transcendentalism in two successive stages. First, I entertain criticisms advanced by Barry Stroud and Richard Rorty against transcendental arguments, arguing that while these criticisms may be telling against certain

contemporary neo-Kantian metaphysicians, they do not apply to the current of neo-Kantian political philosophy that I am defending. Along the way, I make more precise this neo-Kantian position, and argue that Rorty's critique of Kantianism gives us good reasons to prefer Habermas's revision of Kant's meta-ethics over Rawls's. Second, drawing on Apel and Putnam as well as Habermas, I sketch a generalized version of the neo-Kantianism that I argue can withstand Rorty's criticisms. I argue that this generalized position can be understood as a procedural interpretation of ideal limit theory. So the second stage of my argument involves a defence of this procedural interpretation of ideal limit theory, which I argue is a contemporary successor to Kantian transcendentalism. I argue, finally, that Rorty's critique of ideal limit theory, while holding against certain interpretations of ideal limit theory, does not apply to ideal limit theory when given the procedural interpretation I defend.

In the final chapter, I turn to criticisms that approach the political side of the neo-Kantian account of autonomy, focusing on Foucault and a current of the feminist critique of liberalism represented here by Iris Marion Young. Simply stated, Foucault and Young question the assumption in neo-Kantian political theory that the mark of the moral is the universalizable. In this context, we can read Foucault and Young as arguing that it is precisely those interests that do not admit of universalization-- what, in the Kantian tradition, would be called 'pathological'-- that ought to be the object of our moral concern and agenda for

political reform. I then turn to consider the implications of the Young-Foucault critique for the question of the nature of autonomy.

In the conclusion I review the basic elements of my argument, and set the reflections on the nature of autonomy that conclude the fourth chapter in the context of broader concerns in political theory. I end on a programmatic note, offering suggestions concerning the lessons which emerge from the contemporary dialogue in political theory, and suggestions for what I think are the most important avenues to be explored in the dialogue that is to come.

CHAPTER ONE

THE KANTIAN PICTURE

"In some respects", writes Hans Saner, "Kant's political ideas are located at the periphery of his work. It was not until late in the seventh and chiefly in the eighth decade of his life that they were put into cohesively written form, and the space they occupy in his writings is infinitesimal: between 1/2 and 5 percent, depending upon the scope accorded to the word 'political'."¹ The interpretation of Kant that I wish to defend maintains--along with Saner, but, more closely, along with Onora O'Neill--that despite the peripheral status of Kant's writings in the field of political theory, there is a strong political element in the entire critical philosophy.² More strongly, I suggest that there occurs in Kant's writings something of an interface between the epistemic and the political. This interface is hardly peripheral to the critical philosophy; as O'Neill argues, it occurs at the point of the grounding of the authority of reason.

My defense of this interpretation will develop through four sections, the first three concerned primarily with textual security and fidelity, the last with theoretical clarity. In the

¹Hans Saner, Kant's Political Thought (Chicago and London: Chicago University Press, 1973), p. 1.

²Onora O'Neill, "The Public Use of Reason", Political Theory 14:4 (1986), pp. 523-551. Subsequent references to this paper will be inserted into my text.

first section I will consider some general problems with the place of politics in Kant's philosophy, focusing here on Patrick Riley's influential reading of Kant's political philosophy. In the second, I will defend the claim that there is a greater role for politics in Kant than Riley's interpretation admits. This will be done through an examination of Kant's short political essay "What is Enlightenment?" and O'Neill's picture of the place for politics in Kant's philosophy. In the penultimate section I will explore themes highlighted by O'Neill's reading--the subordination of theoretical to practical reason, and the centrality of 'freedom' in the critical philosophy--and take a renewed look at her interpretation. Finally, I will step back from the type of close textual work that characterizes the first three sections of this chapter, and describe the picture that I see emerging from Kant. It is a picture that finds a place for politics among our cognitive and practical activities that, on the one hand, places on a par the epistemic and political conditions that must obtain for the assertions by which we claim knowledge or defend action to be warranted, and, on the other hand, offers a critical space from which to engage in radical political critique. It is not a picture without problems, however. As we will see, it shores up the most difficult problems with Kant's solution to the third Antinomy and his meta-

ethics in general.³

It is a picture, however, that has a hold on many of the seemingly disparate elements of contemporary political theory, as I will spend much of the remainder of the thesis defending.

On the very idea of a Kantian politics

As Patrick Riley suggests, perhaps the central question in Kantian politics is why Kant had a political philosophy at all.⁴ To be sure, Kant did argue that "a constitution allowing *the greatest possible human freedom* in accordance with laws by which *the freedom of each is made to be consistent with that of all others*" was a necessary idea⁵ and an ideal toward which we ought to aspire, along with other political ideals, especially perpetual global peace⁶. However, Kant was also at pains to make clear that as we progress toward these goals "the profit which

³These are the problems that, as we will see in the next chapter, Rawls and Habermas attempt to overcome by 'naturalizing' Kantian autonomy and reformulating what is to count as what Kant called the *pathological* or *heteronomous* determination of the will.

⁴Patrick Riley, Kant's Political Philosophy (New Jersey: Rowman and Littlefield, 1983), p. vii. Subsequent references to this book will be inserted into my text.

⁵Immanuel Kant, The Critique of Pure Reason, trans. Norman Kemp Smith (London: MacMillan, 1929), p. A 316/B 373. Further references to the first Critique will be to this edition and will be inserted into my text in the form of the pagination from the first and second German editions, as indicated by N.K. Smith in his edition.

⁶Cf. "Perpetual Peace: A Philosophical Sketch", in Hans Reiss (ed.), Kant's Political Philosophy (Cambridge: Cambridge University Press, 1970), pp. 93-130.

will accrue to the human race...will not be an ever increasing quantity of *morality* in its attitudes".⁷ No empirical approximation of the kingdom of ends can transcend the merely legal. Such worldly kingdoms can, at best, offer incentives to act *as though* one were acting out of duty with respect to the moral law. In so far as incentives to act for fear of punishment--or otherwise in accordance with the constitution spoken of above--are not, by Kant's account, moral motives, such incentives cannot produce moral actions⁸.

Thus, there seem to be two problems with the very idea of a Kantian politics. (1) Insofar as Kant thought that non-moral motives, because of their contingency, do not command us at all⁹, politics appears somehow superfluous. All we would need, it would seem, is the categorical imperative. And further, if we are bound by duty only to the categorical imperative, then the construction of a just polity cannot be a matter of necessity--as Kant suggested in the quotation above--but rather a matter of prudence, or whimsy. (2) Perhaps we could say that this is too strict. We could understand politics as the handmaiden of morals, appealing to the non-moral motives of the coercive power of the state as positive habituation devices that could bring

⁷"The Contest of Faculties", in Reiss (ed.), p. 187.

⁸Cf. Kant, The Critique of Practical Reason, trans L.W. Beck (New York: Macmillan Publishing Co., 1965), p. 155.

⁹Kant, Groundwork of the Metaphysics of Morals, trans. H. J. Paton (New York: Harper and Row, 1964), p. 86.

about moral action. G.A. Kelly argues, however, that such an appeal would violate Kant's architecture of pure reason, insofar as "a false juncture would [thus] be made between the realms of autonomous and heteronomous causation."¹⁰ To make room for ethics in the Newtonian world, Kant distinguished two types of causation: that which obtained in the phenomenal--or empirical--world, governed by the laws of physics (heteronomous causation) and that which originated in the noumenal--or intelligible--world, governed by the moral law (autonomous causation)¹¹. Kelly's claim is that speaking of non-moral incentives leading one to act morally is to violate this distinction.

Riley argues that, despite these problems, there is room for a Kantian politics. "Kantian public legal justice is purposively related to morality in two ways" (Riley, p. 4), one stronger than the other. In the weaker sense, Riley argues, politics creates the practical conditions for the exercise of the good will "by expanding 'negative' freedom so that one can be 'positively' free, or self-determining through the moral law" (Riley, p. 4). In the stronger sense, politics legally enforces what ought to be, even in the absence of the good will, furthering the moral end of treating persons as ends in themselves. Although I think he is correct in this, there is a more important third place for politics in Kant's architectonic, located by O'Neill's

¹⁰George Armstrong Kelly, Idealism, Politics and History (Cambridge: Cambridge University Press, 1969), pp. 116-117.

¹¹This will be treated in detail below.

interpretation of Kant. I will review Riley's argument, and then, in the next section of this chapter, make my case for this third place.

According to Riley, "Kantian public legal justice is a kind of intersection between the facts of anthropology¹² and the categorical imperative; if there were a kingdom of ends, the kingdoms of earth would vanish" (Riley, p. 3). Against the first of the two objections I listed above, Riley argues that politics is not superfluous, given Kant's description of human nature. In the tradition of social contract theory, Kant feels that, left to our own unbridled devices, human beings would fare little better than in Hobbes' state of nature. Insofar as acting morally means respecting persons as ends in themselves,

if public legal justice sees to it that some moral ends (such as nonmurder) get observed, if not respected, then public legal justice in Kant might be viewed as the partial realization of what would happen if all wills were good. (Riley, pp. 3-4)

As such, the formulation of the categorical imperative that equates the moral with respecting persons as ends in themselves¹³ allows us to understand political action as informed by duty, and at the same time, not force us to collapse legal into moral

¹²Best presented in Kant's observations on the "radical evil" in human nature in the opening sections of Religion Within the Limits of Reason Alone trans. T.M. Greene and H.H. Hudson (New York: Harper & Row, 1960).

¹³and recall that Kant claimed that all three formulations of the categorical imperative are "at bottom, merely so many formulations of precisely the same law." Groundwork, p. 103.

incentives.

Against the objection voiced by Kelly, Riley argues that we can understand politics to be in instrumental service to morality if we say that "public legal justice is instrumental to negative freedom (e.g., freedom from fear), so that persons can be positively free by determining themselves to act from the moral law" (Riley, p. 11).

Mere freedom from fear, taken by itself, would not be moral; but public legal justice, by restricting fear, might diminish an obstacle to moral conduct. Morality, for Kant, is objective; but we can know, from subjective facts of human "pathology," that something like fear may deter us from acting morally. Thus there can be a duty to block--legally--the effect of morality-deflecting fear and appetites. Were that not so, Kantian politics would not be possible at all. (Riley, p. 11)

Kant makes a similar point with respect to the attainment of happiness as the end of actions. While attaining happiness cannot be the end of moral activity, one ought to act to attain happiness *in the service of morals*: "To assure one's happiness is a duty (at least indirectly); for discontent with one's state, in a press of cares and amidst unsatisfied wants, might easily become a great *temptation to the transgression of duty*" (Groundwork, p. 67).¹⁴

So, in Kant, politics is related to morality in two ways. In the weaker sense, the former creates the practical conditions for the exercise of the good will "by expanding 'negative'

¹⁴Kant hints at the same distinction in the Critique of Judgement (Oxford: Oxford University Press, 1952), pp. 47-48.

freedom so that one can be 'positively' free, or self-determining through the moral law" (Riley, p. 4). In the stronger sense, it legally enforces what ought to be, even in the absence of the good will, furthering the moral end of treating persons as ends in themselves.

I will argue that, while Riley solves an important problem in rendering politics compatible with ethics in Kant, politics has a greater role to play in the critical philosophy. Indeed, as O'Neill argues, the entire critical philosophy has a political element. We will get into this interpretation, however, through one of Kant's political writings, "What is Enlightenment?".

The task of enlightenment and the authority of reason

In 1784, Berlinische Monatschrift, a German periodical, solicited from its readers an answer to a most direct question: "What is Enlightenment?", and in November of that year they published a response written by Kant. His answer was simple and unambiguous: "Enlightenment is man's release from his self-incurred immaturity."¹⁵ "Immaturity", Kant goes on to explain, is a person's inability to make use of her or his reason "without direction from another"; this immaturity is "self-incurred" when

¹⁵Kant, "An Answer to the Question: 'What is Enlightenment?'" , in Hans Reiss (ed.), Kant's Political Writings (Cambridge: Cambridge University Press, 1970), p. 54. All further references to this text will be inserted into my text in the form: (WIE, p.__).

its cause lies "not in lack of understanding but in lack of resolution and courage to use it without the guidance of another" (WIE, p. 54). 'Enlightenment' is not so much a state of being or knowing, but rather a process of shrugging off the claims to authority made by anything other than reason itself. As Michel Foucault notes, Kant speaks of 'enlightenment' in an almost entirely negative way, as an 'exit' or a 'way out.' "He characterizes it as a phenomenon, an ongoing process, but he also presents it as a task and an obligation."¹⁶

"For enlightenment of this kind", Kant claims, "all that is needed is *freedom*, freedom, that is, to make public use of reason in all matters (WIE, p. 55). Here Kant introduces his initially puzzling distinction between the public and private uses of reason. The private use of reason--which may remain restricted, and, indeed, whose restriction may in fact further the freedom of the public use of reason (WIE, p. 59)--is "that which a person may make of it in a particular *civil* post or office with which he is entrusted" (WIE, p. 55). That is, for example, a soldier has no right to question an order given to him¹⁷, nor a citizen the right to refuse to pay taxes (WIE, p. 56). By contrast, the total freedom of the public use of reason--"that use which anyone may make of it as a *man of learning* addressing the entire *reading*

¹⁶Michel Foucault, "What is Enlightenment?", in Paul Rabinow (ed.), The Foucault Reader (New York: Pantheon Books, 1984), pp. 34-35.

¹⁷I have suspended use of non-gender specific language here in light of the need for historical fidelity.

public" (WIE, p. 55)--must be defended and upheld. The soldier mentioned above "cannot be reasonably banned from making observations as a man of learning on the errors in the military service, and from submitting these to the public for judgement" (WIE, p. 56). Similarly, while "presumptuous criticisms of taxes, where someone is called upon to pay them, may be punished as an outrage which could lead to general insubordination",

the same citizen does not contravene his civil obligations if, as a learned individual, he publicly voices his thoughts on the impropriety or even injustice of such fiscal measures. (WIE, p. 56)

Similarly, while a member of the clergy is bound to discharge his duties as the church sees fit, he ought to be free to criticize the teachings of the church as a scholar addressing the public.

As Foucault notes, insofar as Kant demands that "reason must be free in its public use, and must be submissive in its private use", his formulation is "term for term, the opposite of what is ordinarily called freedom of conscience."¹⁸ Part of this inversion is accounted for by the fact that Kant is not employing the traditional liberal public/private distinction¹⁹, but a more significant aspect of this inversion lies, as O'Neill argues, deep within Kant's thinking.

The priority which he assigns to the toleration of

¹⁸Foucault, p. 36.

¹⁹Which, as I understand it, would ensure that the private use of reason is free--including here, freedom of worships, freedom of opinion--but would restrict--in respect of the value of tolerance--the demands that can be made on the public order.

public uses of reason has its roots in central Kantian claims about the limits of theoretical reason and the possibility and grounds of practical reason, and the connection of both to the notion of a possible community. (O'Neill, p. 531)

Kant's claims about the limitations of theoretical reason are well known and need not be rehearsed here. His claims about the possibility and grounds of practical reason will be discussed below. The key, at this point, is what O'Neill calls "the connection of both to the notion of a possible community". Kant introduces this point toward the end of the Critique of Pure Reason.

Reason must in all its undertakings subject itself to criticism; should it limit freedom of criticism by any prohibition, it must harm itself, drawing upon itself a damaging suspicion. Nothing is so important through its usefulness, nothing so sacred, that it may be exempted from this searching examination, which knows no respect for persons. *Reason depends on this freedom for its very existence.* For reason has no dictatorial authority; its verdict is always simply the agreement of free citizens, of whom each one must be permitted to express, without let or hindrance, his objections or even his veto. (A 738-739/B 766-767)
[my emphasis]

Thus the authority of reason--the transcendental boundaries of which were established by the first Critique, and upon which the dictatorial force of the categorical imperative is grounded²⁰ (O'Neill, p. 541)--is vindicated only through "the agreement of free citizens", that is, citizens free to make public use of their reason. In O'Neill's words,

The contention is not just that toleration and free discussion will lead to or are necessary for the discoveries of truths...Nor is it that toleration and

²⁰see Kant, Groundwork, pp. 118ff.

free discussion will be politically effective...Such instrumental justifications of toleration all presuppose that we have independent standards of rationality and methods of reaching the truth.²¹ Kant's thought is rather that a degree of toleration must characterize ways of life in which presumed standards of reason and truth can be challenged and so acquire the only sort of vindication of which they are susceptible. The development of reason and of toleration are interdependent...Practices of toleration help constitute reason's authority. (O'Neill, p. 535)

This is the "interface" between the epistemic and the political of which I spoke earlier.

But, it may be contended, why is this a meeting of the *epistemic* and the political, if the type of reasoning with which Kant and O'Neill are concerned is largely moral, or, in Kant's terms, practical? The answer lies in Kant's subordination of theoretical to practical reason (or, in contemporary terms, the subordination of epistemology to ethics). This subordination, in turn, lies in the fact that the Idea²² of freedom, which can only be assumed as a regulative principle by speculative reason--as we will see below--can be granted objective reality by practical reason.

Kant's defence of this latter claim occurs in two steps. In the Critique of Pure Reason he argues that there is at least no

²¹As O'Neill points out, traditional justifications of tolerance within liberal discourse are instrumental; tolerance serves to respect either individuals or reason and truth themselves. (O'Neill, p. 523)

²²that is, "Ideas" in the precise, Kantian sense: not "categories", which are concepts like "causality" that order our possible experience, but rather concepts of pure reason that do not admit of empirical verification: specifically, freedom, the existence of God, and immortality.

contradiction involved in imputing freedom to the action of moral agents. In the last section of the Groundwork and the Critique of Practical Reason, Kant argues that practical reason licences not only the assumption of the free will of moral agents, but proves the objective reality of such a causality. I will turn now to retrace the development of this argument.

Practical and transcendental freedom

'Freedom' is arguably the central concept in the critical philosophy.²³ Kant himself said as much in the second Critique:

The concept of freedom...is the keystone of the whole architecture of the system of pure reason and even of speculative reason. (Practical Reason, p. 3)

But if 'freedom' is the central concept in the architectonic, it is also the most problematic, for while we are 'constrained' by reason to employ it²⁴ (that is, essentially, to impute free will to agents), we cannot properly understand it. This quandary emerges from the central claims of the first Critique. There Kant sought to both answer Humean skepticism and keep a qualified Humean empiricism as a foil against rationalist metaphysics. He did so by asserting that while the boundaries of knowledge were marked by the boundaries of possible experience, experience

²³Riley rejects this claim, and puts *telos* in the place of 'freedom'. But I think he is wrong, precisely because he did not recognize this 'third place' for politics in Kant's philosophy that I am defending on O'Neill's behalf.

²⁴cf. Henry Allison, "Practical and Transcendental Freedom in the Critique of Pure Reason", Kant-studien 773 (1982), pp. 287-288.

itself was possible only through the alignment of sensation under a categorical scheme given by rationality. Our epistemic activities were thus simultaneously bound and freed; while the critical philosophy disallowed departure from the boundaries of experience, restricting our knowledge to things as-they-appear rather than things-in-themselves, within the realm of appearances knowledge was given secure foundations in the a priori conditions of experience. In this way, Kant answered Hume by arguing that if we think--as he imputes to Hume--of our situation as trying to know things-in-themselves, we are correctly led to skepticism. But if we realize that we deal only with things as they appear, then it no longer makes sense to be skeptical about causation, because causality is a precondition of the very appearance of objects within the boundaries of experience; that is, causality is a transcendental precondition of experience²⁵. 'Freedom', or free will, thus seems contrary to the very conditions of experience. But at the same time, freedom (or free will) seems to be a precondition of our *moral* experience. This conflict is one of the antinomies of pure reason that Kant set out to resolve in the Critique of Pure Reason.

Kant's definition of an antinomy (A 340/B 398) is among the more opaque passages of the first Critique. The idea, however, is straightforward: an *antinomy* is a pair of irreconcilable

²⁵For a summary clearer than the account given in the first Critique, see the Critique of Practical Reason, pp. 54-55.

claims²⁶ which cannot be resolved by appeal to experience and whose irresolvability lies in an incorrect understanding of the limits of pure reason. This incorrect understanding lies, basically, in the non-recognition, on the one hand, of the distinction between the phenomenal realm (objects as they appear to us) and the noumenal realm (things-in-themselves), and, on the other hand, of the restriction of theoretical reason to knowledge of the phenomenal world. The Third Antinomy of Pure Reason is the antimony of free will and determinism, i.e.,

"Causality in accordance with laws of nature is not the only causality from which the appearances of the world can one and all be derived. To explain these appearances it is necessary to assume that there is also another causality, that of freedom" vs. "There is no freedom; everything in the world takes place solely in accordance with laws of nature." (A 445/B 473)

Against this formulation of the problem, Kant asks:

Is it a truly disjunctive proposition to say that every effect in the world must arise *either* from nature or from freedom; or must we not rather say that in one and the same event, in different relation, both can be found? (A 536/B 564)

Kant's answer, of course, is "No" to the first option and "Yes" to the second.

The "different relation" of which he speaks in the second formulation of the determinism/free will disjunction above is that which obtains in the phenomenal world, on the one hand, and that which obtains in the noumenal world on the other. Kant's

²⁶concerning, specifically, the ultimate divisibility of substance, the beginning of time and the boundaries of space, free will and determinism, and the existence of an unmoved mover.

solution to the antinomy lies in the application of this distinction: the action of agents is *determined* insofar as it occurs within the order of appearances, which strictly follows the deterministic laws of nature, and *free* insofar as it is determined by reason (A 550/B 578).

It seems to me that Kant is wrestling with two problems here, or that the free will/determinism antinomy presents itself in two different ways to him. The first is an interpretation of the third antinomy in terms of traditional metaphysics, and Kant thinks this solution answers that problem adequately. The second interpretation places the problem in terms of meta-ethics and is more akin to problems of ontological commitment á la Carnap and Quine, and this is the one that runs into problems. The second is the one that interests us here, but let me treat the first briefly as well, to illustrate the distinction I am trying to draw.

A.C. Ewing reads Kant's solution the third antinomy in terms of what I called the metaphysical interpretation. By Ewing's account, "the difficulty for Kant is not that freedom cannot be reconciled with determination by causes, but that it cannot be reconciled with determination by *previous* causes."²⁷ The problem is explained in this argument:

If natural causation is universal, my present acts are

²⁷A.C. Ewing, A Short Commentary on Kant's Critique of Pure Reason (London: Methuen & Co., 1938), p. 228.

always determined entirely by the past, even by my own past.

The past is not in my power now.

Therefore, I am not free now. (Ewing, pp. 228-229)

"But suppose", Ewing continues, that "my acts were determined, but not determined by past events, [then] this difficulty would disappear" (Ewing, p. 229). Insofar as an agent's actions are determined by reason they cannot be spoken of as determined by past events, because reason is extra-empirical, and thus-- according to the Transcendental Aesthetic, wherein Kant argued that space and time are conditions whereby sensation is made possible, just as categories are conditions whereby understanding experience is made possible--outside of space and time.

"Reason", Kant claims, "acts freely; it is not dynamically determined in the chain of natural causes" (A 553/B 581).

Now this account certainly begs questions, and, given that I've summarized a sizable portion of the first Critique very hastily, it has the appearance of a slight-of-hand. I do not have the space to defend this aspect of the solution to the third antinomy. But it is important to include this metaphysical interpretation of the third antinomy because it is important to see that Kant thought this aspect of the antinomy solved. More significantly, it introduces the problem that plagues the resolution of the antinomy as a question of meta-ethics, one that Kant would pursue in the last section of the Groundwork and in the second Critique: how is a *causa noumenon* possible?, i.e., how can we understand the insertion of the dictates of reason

into the current of the deterministic phenomenal world?

What I have called the meta-ethical interpretation of the third antinomy begins with the observation that we *do* have some understanding of the intersection of these two types of causality, or at least our moral discourse seems to involve us in treating agents as though they were susceptible to both forms of causality. (The sense in which the question becomes whether we are thus *committed* to noumenal causality--and thus transcendental freedom--is the sense in which I suggested that the meta-ethical interpretation of the third antinomy is akin to the contemporary question of ontological commitment.) Kant asks us to consider the case of a malicious liar. Let us say we want to investigate the motives for our liar's action, and, on the basis of this investigation, decide to what extent we hold the offender responsible for the action and the consequences of the action. The first part of the investigation is an empirical question, tracing the empirical character of the action to its sources, where we find, e.g., "defective education, bad company, in part also [the] viciousness of a natural disposition insensitive to shame, in levity and thoughtlessness", etc. (A 554/B 582). Although we can explain away the action in terms of natural causes, we nonetheless blame the agent, and thus regard reason "as a cause that irrespective of all the above-mentioned empirical conditions could have determined, and ought to have determined, the agent to act otherwise" (A 555/B 583). The problem, then, is one of giving licence to talking this way, of

making space for moral agency in a world that can only appear to us under the laws of Newtonian determinism.

Kant makes explicit at the end of his discussion of the third antinomy that what he has sought to show is not the reality of freedom, nor even its possibility, but, more modestly, that "this antinomy rests on a sheer illusion, and that causality through freedom is at least *not incompatible with nature*" (A 558/B 596). (The 'sheer illusion' is transcendental realism, the belief that we have contact with things-in-themselves.) Kant thinks that freedom is *at least not incompatible with nature* (determinism) essentially for the reason spelled out by Ewing above: if we picture 'causality through freedom' as a type of vertical causality, emerging from the timeless noumenal realm into the horizontal stream of events in the phenomenal world, there is no strict incompatibility involved in their co-existence. But Kant tried to show more than this, or at least wrote as though he did, insofar as the solution was meant to clear up not only metaphysical but also meta-ethical problems.

Kant distinguishes two types of freedom, transcendental and practical. Freedom in the *transcendental* sense is "the power of beginning a state *spontaneously*" (A 533/B 561). This is the power afforded reason, which, being extra-empirical, is outside time. But by being extra-empirical, freedom in this sense

is a pure transcendental idea, which, in the first place, contains nothing borrowed from experience, and which, secondly, refers to an object that cannot be determined or given in any experience.

(A 533/B 561)

Freedom in the *practical* sense "is the will's independence of coercion through sensuous impulses. For a will is sensuous, in so far as it is *pathologically affected*, i.e., by sensuous motives." However, and this is central to the meta-ethical interpretation of the third antinomy, "sensibility does not necessitate [the human will's] action. There is in man a power of self-determination, independently of any coercion through sensuous impulses" (A 534/B 562). The key point in this passage is Kant's claim that sensibility does not *necessitate* the will's action; practical freedom appears as a means to deal with the type of necessity which is outside of nature: normativity.

That our reason has causality, or that we at least represent it to ourselves as having causality, is evident from the *imperatives* which in all matters of conduct we impose as rules upon our active powers. '*Ought*' expresses a kind of necessity...which is found nowhere else in the whole of nature. The understanding can know in nature only what is, what has been, or what will be. We cannot say that anything in nature *ought to be* other than what in all these time relations it actually is. When we have the course of nature alone in view, '*ought*' has no meaning whatsoever. (A 547/B 575)

Two questions emerge from this explanation of freedom: (1) How are practical and transcendental freedom related? and (2) how does (practical) freedom account for normativity? The answer to the first will lead us to the answer to the second.

(1) In the opening sections of the solution of the third antinomy, Kant argues that "[t]he denial of transcendental freedom must...involve the elimination of practical freedom" (A 534/B 562). He continues,

For practical freedom presupposes that although

something has not happened, it *ought* to have happened, and that its cause, [as found] in the [field of] appearance, is not, therefore, so determining that it excludes a causality of our will--a causality which, independently of those natural causes, and even contrary to our force and influence, can produce something that is determined in the time-order in accordance with empirical laws, and which can therefore begin a series of events *entirely of itself*.

(A 534/B 562)

Recall that Kant characterized practical freedom as "the will's independence from sensuous impulses". Independence from sensuous impulses is independence from the law of causality as it works in the phenomenal realm, insofar as in the phenomenal realm causality is among the conditions of experience. Practical freedom requires transcendental freedom--i.e., "the power to begin a state *spontaneously*"--because for the will to "produce something that is determined in the time order in accordance with empirical laws" (including, most importantly, natural causality), but still be free from determination by sensuous impulses, the will must be able to "insert" itself into the phenomenal realm while being free of the strict determinism that characterizes the world of appearance. (Recall that in what I called the metaphysical interpretation of the third antinomy, Kant argued that, with the application of the noumenal/phenomenal distinction, such "causality through freedom" is at least "not incompatible with nature.") Here is the sense in which I suggested the meta-ethical interpretation of the third antinomy has something in common with the issue of ontological commitment à la Carnap and Quine. Kant is arguing that if we continue to

speak of our moral activities as we do--specifically, if we recognize, in normativity, a type of necessity outside of nature (practical freedom)--then we are committed to the existence of certain transcendental conditions that make these activities possible (transcendental freedom). Kant's argument is not unlike Quine's argument concerning the ontological commitment of theories-- rather than, say, Plato's and Russell's arguments for the existence of universals. Plato and Russell argued that by speaking of things--objects, relations, qualities--we were committed to their existence. Kant is not arguing that simply speaking of freedom commits us to its existence. Rather, Kant's argument for the existence of transcendental freedom seems to be driven by something like Quine's formula that a theory is committed to those entities that must exist for that theory to be true²⁸. Kant argues here that if morality is not to be simply an illusion, we must admit that practical freedom requires there to be transcendental freedom.

2. We have a hint, already, of how practical freedom makes normativity possible: Kant argued in the passage quoted above that "practical freedom presupposes that although something has not happened, it *ought* to have happened." He explained his point succinctly in a passage from the last section of the Groundwork:

²⁸Cf. Quine, "On Carnap's Views On Ontology", The Ways of Paradox and Other Essays, revised ed. (Cambridge: Cambridge University Press, 1976), p. 205, and

Quine, "On What There Is", From a Logical Point of View (Cambridge: Cambridge University Press, 1953), pp. 12-13.

[I]f I were solely a member of the intelligible world, all my actions *would* invariably accord with the autonomy of the will; but because I intuit myself at the same time as a member of the sensible world, they *ought* so to accord. (Groundwork, p. 122)

(Autonomy of the will--a concept I have introduced perhaps belatedly--is, for Kant, autonomy from pathological or sensuous impulses; essentially, practical freedom.) The co-residence Kant is speaking of here can also be understood in terms introduced in the first section of the chapter: we are members both of the kingdom of ends and the kingdoms of the earth. In the sense that, as Riley suggested, if there were really a kingdom of ends, the kingdoms of the earth would vanish, so also if we were solely members of the intelligible world--and thus autonomous, or free from pathological influences--the categorical imperative would cease to be an *imperative*, but would rather be a description of the way persons behave. It is this disjunction, this discordance between our residence in the intelligible and sensible worlds or between the kingdom of ends and the kingdoms of the earth that produces the normative²⁹. According to Kant, this account of normativity re-confirms the point that the existence of practical freedom necessitates the existence of transcendental freedom. In Lewis White Beck's words:

The thought of 'ought' is impossible if all laws are natural laws; the thought of 'ought' implies the thought of a free 'can,' and if pure reason is actually effective in the control of conduct, then there is free causation in the transcendental as well as in the

²⁹As we will see in chapter three, this meta-ethics is paralleled by what may be called "ideal limit theory"; here the kingdom of ends is replaced by the end of inquiry.

practical sense.³⁰

We are now in a position to understand Kant's subordination of theoretical to practical reason, the theme emerging from O'Neill's reading of Kant that prompted this investigation into Kant's account of freedom.

It should be made clear, first of all, that it would be incorrect to think of theoretical and practical reason as different faculties. If a critique of practical reason is to be complete, Kant claims, we should be able "to show the unity of practical and theoretical reason in a common principle, since in the end there can only be one and the same reason, which must be differentiated solely in its application" (Groundwork, p. 59). (This common principle is freedom.) 'Practical reason' is pure reason (that is, reason concerned with matters a priori) behaving practically. The practical, Kant says, is that which is possible through freedom. Freedom, as I explained above, is what gives us normativity; so, if pure reason is to be practical, it must be normative. For pure reason to be practical it must "of itself and independently of everything empirical [be able to] determine the will" (Practical Reason, p. 43), and, as L.W. Beck suggests in the introduction to his translation of the second Critique, not act "merely as 'the servant of the passions' (Hume), i.e., in

³⁰Lewis White Beck, A Commentary on Kant's Critique of Practical Reason (Chicago: Chicago University Press, 1960), p. 189.

connection with other, non-rational components of personality."³¹

Now, the desire to have pure reason act practically is not a qualified one; it is not the case, according to Kant, that meta-ethics will simply *benefit* by enjoying such a grounding. It is rather the case that, as Beck points out, pure reason *must* be practical if morality is not an illusion.³² To make sense of this, we must recall Kant's account of normativity. The normative arises, in Kant, as a function of the tension between what would obtain if we lived in the kingdom of ends and the fact that we live in the kingdoms of the earth. Recall, as well, that the 'ought' of moral imperatives is a matter of necessity, rather than a matter of contingency--or the pathological determination of the will. For normativity to be a matter of necessity, the 'ought' of moral imperative must hold *no matter what*; it must hold independently of any particular interest: it must hold, that is, a priori. As such, if morality is not an illusion--that is, if moral imperatives are not simply pieces of prudential advice, maxims that serve contingent self-interest--the moral must be knowable a priori.

Kant gave us such an a priori ethics with the formulations of the categorical imperative in the Groundwork. But there is an epistemic problem with the moral law that lies in the fact that

³¹L.W. Beck, "Translator's Introduction", to Critique of Practical Reason, p. xii.

³²Beck, "Translator's Introduction", Practical Reason, p. xii.

"categorical imperatives are possible because the Idea of freedom makes me a member of an intelligible world" (Groundwork, p. 122). The Idea of freedom is beyond the scope of theoretical or speculative reason, which the Transcendental Analytic of the first Critique limited to the world of appearances, wherein the law of strict causality rules. While Kant's solution to the third antinomy argued that causality through freedom is at least not incompatible with the determined order of experience, such causality cannot be *known* but can only be *thought* by speculative reason, whose boundaries of legitimacy are co-extensive with the boundaries of possible experience. "With the pure practical faculty of reason," however, "the reality of transcendental freedom is also confirmed" (Practical Reason, p. 3). The crucial point is that

Reason is not hereby extended, however, in its theoretical knowledge; the only thing is that the possibility [of freedom], which was heretofore a problem, now becomes an assertion, and the practical use of reason is thus connected with the elements of theoretical reason. (Practical Reason, p. 4)

While the need to solve the third antinomy 'constrains' speculative reason to *think* of freedom, Kant argues that pure reason in its practical employment gives objective reality to freedom.

The argument for this latter claim can be reconstructed from what has been discussed already.

- P₁: Pure reason must be practical, if morality is not an illusion.
- P₂: For pure reason to be practical, it must be able to determine the will in and of itself.

P₃: For pure reason to be able to determine the will, the will must be able to act independently of sensuous impulses.

P₄: The will can act independently of sensuous impulses; that is, there is practical freedom.

Therefore, pure reason can be practical.

P₅: If there is practical freedom, there must be transcendental freedom.

Therefore, insofar as pure reason can be practical, transcendental freedom exists.

The weakest step in this argument appears to be P₄. Kant provides no argument for this claim, and appears to take it to be an indisputable fact.

[The fact of] practical freedom can be proved through experience. For the human will is not determined by that alone which stimulates, that is, immediately affects the senses; we have the power to overcome the impressions on our faculty of sensuous desire.
(A 802/B 830)

By Kant's account, however, the apparent weakness of P₄ demonstrates the superiority of practical over theoretical reason³³. The 'weakness' of this claim lies in the possibility of an appearance/reality disjunction; it may be the case that our experience of practical freedom is an illusion. In response to this, Kant argues that

[w]hether reason is not, in the actions through which it prescribes laws, itself again determined by other influences, and whether that which, in relation to sensuous impulses, is entitled freedom, may not, in relation to higher and more remote operating causes, be nature [i.e., determined] again, is a question which in

³³Let me repeat that "practical over theoretical reason" is elliptical for "pure reason in its practical employment over pure reason in its theoretical employment."

the practical field does not concern us, since we are demanding of reason nothing but the *rule* of conduct; it is a merely speculative question, which we can leave aside so long as we are considering what ought or ought not to be done. (A 803/B 831)

So practical reason regards with indifference the raising of the appearance/reality disjunction against the 'fact' of practical freedom. While theoretical reason may be driven to raise these doubts--and thus re-create the third antinomy--practical reason need not and cannot, for, as Beck puts it,

in the practical [realm], the Idea of freedom is constitutive of the experience to which it applies, for the experience is of what ought to be (as defined by the Idea) and not of what happens to exist independently of it.³⁴

It is in this sense that practical reason can affirm the reality of transcendental freedom³⁵, while theoretical reason is no more than 'constrained' to think it. This is why Kant subordinates theoretical to practical reason.

So we are back, finally, to O'Neill's reading of Kant. At the end of the last section I was defending, with O'Neill's help, the claim that there occurs in Kant something of an interface between the epistemic and the political. This interpretation turned on Kant's claim that "reason has no dictatorial authority; its verdict is always simply the agreement of free citizens" (A 738/B 766). We were led to consider the

³⁴L.W. Beck, A Commentary on Kant's Critique of Practical Reason, p. 48.

³⁵which, recall, follows from practical freedom

relationship between practical and theoretical reason because it seemed that here Kant was speaking more of the former; it seems that the first Critique established the foundations of science independently of anything like consensual vindication. My point and O'Neill's point (as I read her) is that this *is* a meeting of the epistemic and the political insofar as theoretical reason is subordinated to practical reason; or, in contemporary terms, epistemology is subordinated to ethics.

But how far have we gotten with establishing this claim? As O'Neill points out, "[e]ven if we were to accept these arguments [for the subordination of theoretical to practical reason], we would still be in the dark about the grounds of practical reason" (O'Neill, p. 525).

It is not enough to say that *if* we reason theoretically then we must also be able to or be committed to reasoning practically. It seems rather that we should be able to see *why* the standards we recognize as rational in practical matters are these standards, and not others. Yet how can this demand ever be met? We appear to be faced with a familiar dilemma. If the standards of practical reasoning are fundamental to all human reasoning, then any vindication of these standards is either circular (since it uses those very standards) or a failure (since it is not a vindication in terms of the standards said to be fundamental). What then can be said on behalf of the standards of practical reasoning? (O'Neill, p. 525)

Much of the answer to this question is contained in what has been discussed in the second and third sections of this chapter.

Kant's vindication of the authority of reason lies somewhere between the transcendental and the empirical. It is this tension that makes for what I have called the epistemic/political

interface; were it purely transcendental, the political would be secondary, and were it purely empirical, the epistemic would be secondary. O'Neill, I believe, misses this point when she claims that "[r]eason, on this account [i.e., the account based on Kant's claim that the authority of reason extends as far as the authority granted by 'the agreement of free citizens'] has no transcendental foundations but is based rather on agreement of a certain sort" (O'Neill, p. 534). The key phrase here is *agreement of a certain sort*; a standard for agreement is presupposed, and the criteria for this standard are transcendental. O'Neill makes *this* point clearly:

Mere agreement, were it possible, would have no authority.³⁶ What makes agreement of a certain sort authoritative is that it is agreement based on principles that meet their own criticism. The principles of reason vindicate their authority by their stamina when used recursively. (O'Neill, p. 534)

As we will see in chapter three, this 'recursive stamina' is

³⁶Kant backs up this point in the Critique of Practical Reason. "[U]niversality of assent does not prove the objective validity of a judgement, i.e., its validity as knowledge, but only calls attention to the fact that, even if sometimes that which is universally assented to is also correct, this is no proof of its agreement with the object; it is rather the case that only objective validity affords the ground of a necessary universal agreement." (Practical Reason, p. 13) Kant's phrasing here is perhaps unfortunate; it is at least misleading. He seems to suggest that a judgement is valid iff it "agrees with its object", suggesting some kind of correspondence theory of truth. Kant's insistence on the restriction of knowledge to the phenomenal realm seems--and is--incompatible with such a conception of truth (as least as it is usually understood). But he clarifies himself by turning, in the next sentence, to speaking of "objective validity". I suggest that the distinction Kant is drawing here might be translated in more contemporary terms into the distinction between 'mere' agreement and rational consensus.

perhaps the central mark of the transcendental status of a standard or criterion. The principles of reason exhibit such recursive stamina insofar as "they meet their own criticism." What O'Neill, on behalf of Kant, is suggesting here is that the political value of toleration--as the value that underscores the "agreement of free citizens" as a criterion for that which is to have authority--does not suffer the same fate as, say, Popper's criterion of 'falsification' when applied to itself.³⁷ Free citizens would not agree that it is warranted to assent to propositions agreed to under conditions of coercion.

So the vindication of the authority of reason lies somewhere between the transcendental and the empirical--and thus points to an epistemic/political interface--insofar as such a vindication is possible only through consensus (which is empirical), but only through consensus that is *a priori* rational (and these standards of rationality are transcendental). Here is the core of the 'Kantian picture' that I am trying to draw: the preconditions of such rational consensus is freedom, *both* practical/transcendental *and* political. 'Freedom' in Kant has a dual role; as G.A. Kelly

³⁷Popper offered as a criterion demarcating science from pseudo-science that (properly) scientific theories admit of falsification. It is not clear how this criteria *itself* admits of such vindication; i.e., should Popper not falsify himself?

suggests, it is both "an a priori concept and a mission."³⁸ It is an a priori concept in the sense articulated in the Critiques; it is a mission in the sense described in "What Is Enlightenment?". When asked "What is Enlightenment?", Kant responded that it is "man's release from his self-incurred immaturity" and that it is our duty to strive toward this liberation, which could be won only by securing autonomy from others and from tradition. For Kant, rationality provides the mechanisms for this escape from contingency: as noumenal beings, we already travel in the realm of freedom. It is from this position of autonomy and freedom that we can ground our critique of the institutions that deny us the free use of our reason, and it is toward the securing of political preconditions of autonomy and freedom that our critique is to be directed. My suggestion is that this basic picture still has a hold on contemporary political theory, and that many writers today are either proposing solutions to certain problems in the details of Kant's architectonic, or trying to convince us that we can do without this elaborate scheme of justification.

I will turn now to review this picture, and consider the problems with Kant that it brings into relief; problems that, as we will see, in their contemporary form make up much of the substance of the debates in contemporary political theory.

³⁸Kelly, p. 123.

The Kantian picture reviewed

As I indicated in the introduction and first section of this chapter, finding a place for politics in Kant poses a number of hermeneutical problems. These problems derive largely from the apparent peripheral status of Kant's writings in politics within his corpus and from the difficult status any kind of empirical enforcement of behaviour has in light of Kant's ethics. These problems have led writers in the recent revival of interest in Kant's politics to propose different algorithms for generating a consistent politics from Kant. Many writers have taken recourse to either of the two sections of the Critique of Judgment. Riley's book argues that Kant's politics can be given a place in the rest of his philosophy through the common concept of teleology, which is given its most sustained investigation in the "Critique of Teleological Judgement." Steven Delue argues that with the help of the analysis of the beautiful in the first half of the third Critique--"The Critique of Aesthetic Judgement"--we can understand 'respect for persons' as the central concept in Kant's politics and a principle unifying Kant's politics with the critical philosophy.³⁹

My interpretation, by contrast, resists this contemporary

³⁹Steven M. Delue, "Kant's Politics as an Expression of the Need for His Aesthetics", Political Theory 13:3 (1985), pp. 409-429.

The turn to the Critique of Judgment occurs elsewhere, most notably with Hannah Arendt, in her Lectures on Kant's Political Philosophy, Ronald Beiner (ed.), (Chicago: Chicago University Press, 1982), and in Beiner's interpretive essay included in that volume.

turning to the third Critique. My claim is that Kant's politics is most intimately related to his investigations into the authority of reason--that is, to his epistemology and meta-ethics. As such, my reading concentrates on sections of the first Critique, the last section of the Groundwork, the Critique of Practical Reason, and the briefest of Kant's political essays, but the one most concerned with the relationship of reason and authority, "What Is Enlightenment?". The Kantian picture, as I have drawn it, derives largely from an interpretation of the relationship between "What Is Enlightenment?" and the critical philosophy, an interpretation given lucid expression by Foucault.

[In WIE,] Kant in fact describes Enlightenment as the moment when humanity is going to put its own reason to use, without subjecting itself to any authority; but it is precisely at this moment that the critique is necessary, since its role is that of defining the conditions under which the use of reason is legitimate in order to determine what can be known, what must be done, and what may be hoped⁴⁰. Illegitimate uses of reason are what give rise to dogmatism and heteronomy, along with illusion; on the other hand, it is when the legitimate use of reason has been clearly defined in its principles that its autonomy can be assured. The critique is, in a sense, the handbook of reason that has grown up in Enlightenment; and, conversely, the Enlightenment is the age of the critique.⁴¹

What Foucault is pointing to in this passage is that the account of political autonomy (the free use of public reason) given in

⁴⁰"All the interests of my reason, speculative as well as practical, combine in the three following questions:

1. What can I know?
2. What ought I to do?
3. What may I hope?"

Kant, Critique of Pure Reason (A 804-805/B 832-833).

⁴¹Foucault, pp. 37-38.

"What is Enlightenment?" and the account of meta-ethical autonomy (practical and transcendental freedom) given in the critical philosophy are related in something of a symbiotic relationship. This relationship is the one I described at the end of the last section, where I argued that Kant's vindication of the authority of reason lies somewhere between the transcendental and empirical. In contemporary terms, we may say that Kant argues that we are warranted in assenting to or acting on propositions which enjoy vindication through 'rational consensus', and rational consensus has both epistemic and political elements, neither of which can be reduced to the other.

As I suggested in the introduction to this chapter, the Kantian picture, as well as offering an account of the relationship between the epistemic and the political, offers a critical space from which to engage in radical political critique. This space is the space between the kingdom of ends and the kingdoms of the earth. The dual nature of autonomy and freedom in Kant re-appears: insofar as we are, in some sense, free and capable of acting in moral autonomy, we can recognize the political institutions and practices that hamper the empirical realization of the conditions of autonomy (i.e., the free use of reason). It is toward the dismantling of these oppressive institutions and the reform of these oppressive practices that our critique is to be directed. Autonomy, in this sense, appears at the beginning and the end of Kantian philosophy.

These two sides of the Kantian picture--the one creates what I have called an epistemic/political interface, and the other which offers a foundation for radical political critique--are both underscored by Kant's meta-ethics in general and his account of normativity in particular. This meta-ethics, in turn, draws on the noumenal/phenomenal distinction. The normative appears out of the tension given by our co-residence in the noumenal and phenomenal realms, or the determination of our behaviour by both rational and pathological motives. If we were purely noumenal beings, that is, if our behaviour were free from influence by pathological incentives, the categorical imperative would cease to be an *imperative*, and would simply be a description of how persons behave.

But if the noumenal/phenomenal distinction is thus the central concept in Kant's meta-ethics, it is also the most problematic. This distinction had to be invoked to make ethics possible, but Kant never gave a satisfactory account of how the interaction between the noumenal and phenomenal was possible; how, that is, the timeless dictates of the categorical imperative insert themselves into the deterministic world of appearance. The obvious answer, that the relationship is causal, seems ruled out insofar as causality is one of a categories of understanding, and, as such, meaningful only in terms of possible experience. Kant did argue in the second Critique, however, that causality could be applied to the noumenal realm. This, Kant claims, is part of the "enigma" of the critical philosophy,

which lies in the fact that we must renounce the objective reality of the supersensible use of the categories in speculation and yet can attribute this reality to them in respect to the objects of pure practical reason. (Practical Reason, p. 5.)

Kant argued later in the second critique that, insofar as its origin is independent from all sensuous origins, the use of the concept of causality is not restricted to phenomena, and can be applied to the moral law insofar as it is thus used practically. (Practical Reason, pp. 57-58).

I do not have the space to begin to evaluate Kant's response to this problem. It strikes me that it will not do--he seems here to be undermining what I take to be the central insight of the first critique: that while the boundaries of knowledge are co-extensive with the boundaries of possible experience, experience itself is made possible only by its interpretation under a scheme of concepts. This picture requires that the restrictions Kant placed on these concepts (including causality), i.e., that they have meaning only in terms of possible experience, stay in place. The explanation of *causa noumenon* in the second Critique comes into direct conflict with this.

Perhaps there is a way to rescue Kant from this problem. Rather than pursue it, however, I would like to address this question from another angle (one that leads us into the contemporary relevance of this issue): what was it that led Kant to appeal to the noumenal/phenomenal distinction (and thus to the concept of *causa noumenon*) in his meta-ethics in the first place? We saw above that the third antinomy, given what I called the

metaphysical interpretation, required this appeal to make room for ethics in a world that appears to us under the laws of Newtonian determinism. Approached from the meta-ethical side, however, the noumenal/phenomenal distinction offers to Kant a means by which he could draw a strict demarcation between what he took to be the pathological and what he took to be rational. What got in the way of morality, for Kant, was the sensuous determination of the will, which could only be contingent, and thus could be the determining ground of no more than prudential advice. For reason to have authority, its appeal must be universal; as such, appeal made on pathological and contingent grounds could offer no imperatives.

Most contemporary political theorists would not regard the source of pathology to be in the sensuous, but nonetheless would regard it to be in the contingent, with the contingent being typically reformulated as that which is in the interests of some, but not all. Two central questions emerge from this move. (1) Can we reformulate Kant's account of the pathological, and keep his account of the normative (and with it, his defence of radical political critique)? As we will see, this revision is the essential link among contemporary neo-Kantians, who attempt, in one way or another, to 'naturalize' the Kantian account of autonomy (that is, in Kant's terms, keep it all in the phenomenal realm). The second question emerges from this move: (2) can this naturalization, while escaping the positing of a noumenal realm, also avoid taking recourse to a transcendental vindication of the

authority of reason? Much of the rest of my thesis will be an explorations of these questions through an examination of the answers given to them by the dominant currents in contemporary political theory.

I will begin this investigation in the next chapter with an examination of John Rawls's and Jürgen Habermas's answers to these questions.

CHAPTER TWO**KANTIAN AUTONOMY NATURALIZED**

At first glance, John Rawls and Jürgen Habermas may seem to have little in common, either methodologically or ideologically. Rawls's work is paradigmatic analytic political philosophy, written in the context of a tradition that understands political philosophy to be a sub-discipline of ethics. His Theory of Justice offers a vindication of welfare-minded liberalism. Habermas's work, by contrast, seems a paradigm of continental philosophy, written in the context of the inter-disciplinary tradition of the Frankfurt School. Habermas's writings offer a vindication of democratic socialism.

I will argue, however, that both their programs have at their core two critical revisions of Kant that are strongly analogous.

(1) Both Rawls and Habermas offer procedural re-interpretations of Kant's meta-ethics. In the place of noumenal autonomy, Rawls and Habermas recast moral autonomy in terms of the impartiality afforded to groups of rational persons deliberating under constraints which exclude the morally contingent: the "veil of ignorance" and the "ideal speech situation", respectively. Such a re-interpretation of Kant's meta-ethics amounts to a 'naturalization' of Kant's account of

autonomy¹.

(2) Both Rawls and Habermas, in articulating the conditions of the 'veil of ignorance' and the 'ideal speech situation' re-interpret what Kant called the pathological. Both accept that the pathological is the contingent², but reject that it is

¹I should explain my use of 'naturalization' here. Naturalism takes many forms: programs of reducing epistemic factors to psychological or sociological factors; programs of rendering the methodology of the philosophy of science continuous with the methodology of science; programs of reducing value terms to moral facts, and then in turn to natural facts; to name a few. What all these programs have in common is a desire to level normative hierarchies, and in doing so, de-proliferate theoretical entities or level of abstraction from a given normative framework.

For example, a naturalized philosophy of science aims to knock the a priori scaffolding out from under methodology understood as a meta-level system of constraint that draws the boundaries of legitimacy for scientific practice. Methodological questions are put on an epistemic par with the substantive questions of natural science. By Larry Laudan's account, this involves translating the 'categorical imperatives' of methodology (e.g., avoid ad hoc additions to hypotheses) into 'hypothetical imperatives' (e.g., if you want greater empirical adequacy in the long run, avoid ad hoc additions to hypotheses), and translating these in turn to hypotheses that admit of empirical testing (e.g., if I avoid ad hoc additions to hypotheses, will I achieve greater empirical adequacy in the long run?). In this way, methodological questions face the same tribunal as substantive scientific questions.

This results in abandoning a level of abstraction from the normative framework of scientific methodology, i.e., the a priori. Similarly, by offering interpretations of Kant's meta-ethics that replaces the autonomy Kant located in the noumenal realm with autonomy secured by impartial reasoning, Rawls and Habermas rescue the possibility of ethics from the need to posit the noumenal realm. 'Autonomy' is thus understood, in Kant's terms, to be only within the phenomenal realm (and not also in the noumenal realm). In the sense that our normative framework is thus made more austere, Rawls and Habermas have 'naturalized' Kantian autonomy.

²Or, in more strictly Kantian terms, that the will is pathologically determined if it's grounds of determination are contingent.

coextensive with the sensuous. For Rawls the pathological becomes, essentially, knowledge of your position in society. For Habermas, the pathological becomes the ideological.

This chapter is divided into two sections, the first primarily exegetic, the second critical. In the first, I describe the veil of ignorance and the ideal speech situation, and explain how they are grounded in the two revisions listed above. The section on Rawls will be shorter than that on Habermas, due to the relative familiarity of the former and the relative complexity of the latter. In the second, I evaluate these revisions in terms of two concerns:

(1) How far may the Kantian interpretation of Rawls and Habermas be pushed? That is, have I overstated their debt to Kant (and, as such, their common ground)?

(2) With Rawls's and Habermas's revision in hand, I return to the questions I asked of contemporary neo-Kantianism at the end of the first chapter: (a) Can we reformulate Kant's account of the pathological, and keep his account of the normative (and with it, his defence of radical political critique)?, and (b) Can naturalizing Kantian autonomy, while escaping the positing of a noumenal realm, also avoid taking recourse to a transcendental vindication of the authority of reason?

Throughout the chapter, my primary concern is to defend two claims: (1) that Kant's meta-ethics in general and his account of autonomy in particular have relevance for the concerns being

addressed in contemporary political theory and are represented in the neo-Kantianism of Rawls and Habermas; and (2) that Kant's meta-ethics in general and account of autonomy in particular can be recast in light of contemporary concerns (and are so re-cast by Rawls and Habermas) and retain the formal structure that I argued (in chapter 1) served as a grounding, for Kant, for radical political critique.

The veil of ignorance and the ideal speech situation

Rawls's theory of justice can be separated into two levels: his substantive principles of justice--the principle of equality and the difference principle--and the mechanism of justification he provides for them--the original position³. My concern here is

³Rawls, I suspect, would be unhappy with the ease with which I have stated this distinction, as his point seems to be that if we accept the intuitive appeal of the original position, then his two principles follow with almost deductive security. It is not my point to challenge this--though many of his critics have (both by suggesting that they do not follow, or, if they do, then the argument is circular)--but to isolate, for the purposes of this discussion, the motivations behind the construction of the original position.

with the latter⁴. Rawls's theory is grounded in an intuition captured by the phrase "justice as fairness": the just is what would be agreed upon by rational agents placed under certain constraints which insure that the agreements reached are fair, "fairness" being, essentially, "impartiality". Justice as fairness (hereafter JAF) is thus cast as a (non-Hobbsian, non-Nozickian) social contract theory. Social contract theories, Rawls argues, suggest the importance of plurality and publicity to questions of justice. Most importantly, Rawls suggests, casting JAF in terms of a social contract has the merit of conveying the idea that "principles of justice may be conceived as principles that would be chosen by rational persons, and that in this way conceptions of justice may be explained and justified."⁵ The theory of justice is part, perhaps the most

⁴There is a further reason for limiting my consideration of Rawls to the meta-level. One of Habermas's differences with Rawls lies in Habermas's conviction that a theory of justice ought to remain at the level of spelling out the procedural conditions whereby subjects may rationally deliberate questions of the just. It is not that I wish to overlook their differences in an attempt to further my reconstruction of Rawls and Habermas by emphasizing their commonality in the Kantian tradition. It is, rather, that this latter similarity is limited to the meta-theoretical realm; Rawls and Habermas are in basic agreement only as far as they are Kantians. That is, Rawls's principles of justice are not particularly Kantian (and neither are Habermas's more substantive critiques of advanced capitalism).

⁵Rawls fills out the sense in which this is a social contract theory, suggesting that the original position can be considered analogous to the state of nature in traditional social contract theory [John Rawls, A Theory of Justice (Cambridge: Belknap Press, 1971), p. 11; hereafter references to this book will be made in my text in the form: (TOJ, p. __)]. As with the state of nature (in many interpretations), the original position is hypothetical, and works as a kind of counterfactual justification. "[A] society satisfying the principles of justice

significant part, of the theory of rational choice" (TOJ, p. 13). Rational choice is secured by the deprivation of knowledge. Agents are placed behind the "veil of ignorance", which is constructed to exclude matters of contingency: "no one knows his place in society, his class position or social status, nor does anyone know his fortune in the distribution of natural assets and abilities, his intelligence, strength and the like. I shall even assume that the parties do not know their conception of the good or their special psychological propensities" (TOJ, p.12).⁶ They do know, however, "the general facts about human society", which include a grasp of politics, economics and psychology (TOJ, p.137), and they are allowed to assume a preference for more basic social goods rather than fewer (TOJ, p. 142). In sum,

The idea of the original position is to set up a fair procedure so that any principles agreed to will be just. The aim is to use the notion of pure procedural justice⁷ as a basis of theory. Somehow we must nullify the effects of specific contingencies which put men at odds and tempt them to exploit social and natural circumstances to their own advantage. Now in order to do this I assume that the parties are situated behind a veil of ignorance. (TOJ, p. 136)

So it is not that rational agents under the veil of ignorance are

as fairness comes as close as a society can to being a voluntary scheme, for it meets the principles which free and equal persons would assent to under circumstances that are fair" (TOJ, p. 13).

⁶See TOJ, p. 137 for a richer description of the conditions of the veil of ignorance.

⁷"[P]ure procedural justice obtains when there is no independent criterion for the right result [of the inquiry]: instead there is a correct or fair procedure such that the outcome is likewise correct and fair, whatever it is, provided that the procedure has been properly followed." (TOJ, p. 86)

thus allowed to *recognize* the just; this would open Rawls up to the Euthyphro critique⁸ and make his account of rational choice debilitatingly circular. Rather, the just *is* what would be chosen by rational agents drawing a social contract in the original position. What *is* chosen tells us how we, as citizens, *ought* to behave⁹.

Rawls makes explicit links between his theory and Kantian ethics, suggesting that "the original position may be viewed...as a procedural interpretation of Kant's conception of autonomy and the categorical imperative" (TOJ, p. 256). The veil of ignorance in effect allows the agents in the original position to act in the freedom that Kant argued was prerequisite for rational moral decisions, insofar as the barring of what Rawls calls "arbitrary contingencies" is analogous to barring what Kant called the "pathological". Rawls thus suggests that "we think of the original position as the point of view from which noumenal selves

⁸Euthyphro offers as a definition of holiness, "what is pleasing to the gods is holy, and what is not pleasing to them is unholy." (6e-7a) Socrates asks "Is what is holy holy because the gods approve it, or do they approve it because it is holy?" (10a). If it is the latter, then Euthyphro's definition just begs the question, because if the gods approve of it because it is holy, then they simply recognize its holiness, i.e., there must be an independent criterion. Similarly, it can't be the case that the agents in the original position simply recognize what is just. In Rawls' terms, if this were the case, his theory of rational choice would not be following pure procedural justice.

⁹The nature of the is-ought inference will be explored in more detail in the second section of this chapter, where I will take up the question of whether Rawls (and Habermas), while reformulating Kant's account of the pathological, keep Kant's account of the normative.

see the world" (TOJ, p. 255). The principles of justice chosen in the original position are thus categorical (rather than hypothetical) imperatives in Kant's sense; categorical imperatives, according to Rawls, are "principle[s] of conduct that [apply] to a person in virtue of his nature as a free and equal being" (TOJ, p. 253). Kant's exclusion of any particular conception of the good from the matters of ethics is paralleled by Rawls's exclusion of any particular conception of the good from among the items of knowledge allowed agents behind the veil. The universality of the principles of justice are thus secured in a manner analogous to Kant's categorical imperative.

To act from the principles of justice is to act from categorical imperatives in the sense that they apply to us whatever in principle our aims are. This simply reflects the fact that no such contingencies appear as premises in their derivation. (TOJ, p. 253)

So, for Rawls, as for Kant, universalizability is the criterion by which a principle may be binding. A principle may be universalized, and thus hold categorically, if it applies to all agents "whatever in principle [their] aims are." So as Kant sought to exclude, from the moral, the contingent--which he identified with sensuous impulses and material goals--Rawls's veil of ignorance acts to level out contingent differences among the rational contractors.

As well as recasting the pathological in terms of the intrusion of morally arbitrary knowledge in the process of moral deliberation, Rawls recasts Kant's understanding of autonomy in

procedural terms¹⁰. Agents are acting autonomously, for Rawls, when they are afforded the impartiality given by the constraints of the veil of ignorance to deliberate over questions of justice. But with Rawls, autonomy is 'naturalized' into the phenomenal realm. The original position serves as a rational ideal which constrains our moral deliberation (much as theoretical reason was 'constrained' to posit freedom), but is no more than a regulative ideal (that is, it need not be given 'objective reality' as Kant felt practical reason granted freedom). According to Rawls, pure procedural justice¹¹ allows us to formulate universally binding principles of justice--and thus makes morals possible--and the conditions for this pure procedure are contained within the criteria of impartial reasoning--and thus the possibility of morals is saved from the need to posit the noumenal realm.

Like Rawls's "veil of ignorance," Habermas's "ideal speech situation" is a kind of procedural interpretation of Kant's meta-

¹⁰An important question to ask Rawls concerning the status of his theory as a *procedural* reinterpretation of Kant is why, exactly, does Rawls have to cast his theory in terms of a social contract. If agents behind the veil of ignorance are in an equal position with respect to knowledge and interest, why wouldn't they immediately agree on all matters? In this sense, Rawls seems to be sliding toward an ideal observer justification of his principles of justice.

The general problems with creating rational deliberators by abstracting contingencies to the point of establishing equality will be taken up in detail in the fourth chapter of this thesis, wherein I will turn to consider what I understand to be the most formidable challenges to the neo-Kantian tradition.

¹¹i.e., deliberation in the absence of an independent criterion for the correct result of a given inquiry.

ethics. Unlike Rawls, however, Habermas does not seek to articulate the conditions of rational arbitration to ground a particular normative ethics, as Rawls does in justifying his substantive theory of justice by reference to its favourable status among agents behind the veil of ignorance. Rather, Habermas develops the concept of the 'ideal speech situation' within the context of articulating a foundation for the critique of ideology. The search for such a foundation is the motivation behind much of Habermas's work¹². It is at the point of developing this foundation that Habermas engages in his revision of Kant, which, as it was with Rawls, centers on a re-interpretation of what Kant considered to be the pathological motivation of the will.

Habermas, like Rawls, accepts Kant's identification of the pathological with the contingent, but rejects Kant's identification of the latter with material interest. For Habermas, the pathological is, roughly speaking, the ideological. There are two aspects of 'ideology' as Habermas understands it that constitute its role as the pathological.

(1) Like Kant (and Rawls), Habermas is a moral cognitivist, and takes universality to be the mark of the moral. But unlike Kant, for Habermas the unsuitability of any particular interest for universal legislation lies not in the fact that it is an

¹²I should note that this statement of motivation is not something Habermas makes explicitly, but is rather an element of my reconstruction of his writings.

interest (and thus born of sensuous impulses), but in the fact that it is non-generalizable¹³. Now, since Marx and Engels, the ideological has been linked to contingent interests. I take the kernel of Marx and Engels' conception of ideology to be captured in the following observations from The German Ideology:

[E]ach new class which puts itself in the place of one ruling before it, is compelled, merely in order to carry through its aim, to represent its interests as the common interest of all its members of society, that is, expressed in ideal form: it has to give its ideas a form of universality, and represent them as the only rational, universally valid ones.¹⁴

So, basically, according to Marx and Engels, the ideological is that which is in the interests of some, but is represented as being in the interests of all; or that which is contingent (or of contingent interest) but is represented as universal. While Habermas offers substantial revisions on the classical Marxist conception of ideology--the most important of which will be taken up below--it seems the kernel of this classical conception remains intact in his thinking. In this (classical) sense, ideology plays the role in Habermas's meta-ethics that the pathological determination of the will did in Kant's meta-ethics insofar as an ideological interest--being particular--is an interest that does not admit of universalization (but is

¹³Cf. Thomas McCarthy, The Critical Theory of Jürgen Habermas (Cambridge: MIT Press, 1978), pp. 328. This point will be spelled out in greater detail below.

¹⁴Karl Marx and Frederick Engels, The German Ideology, from T. Borodilina (ed.), On Historical Materialism (Moscow: Progress Publishers, 1972), p. 45.

presented as though it does). Analogously, Kant excluded sensuous impulses because he felt that the attainment of their objects could not be universalized as a goal.

(2) While, as I suggested, the core of the classical Marxist notion of ideology remains relevant to Habermas's thinking, he does argue that advanced capitalism differs so much from the 'liberal capitalism' of Marx's day that 'ideology' must be revised if it is to remain a relevant critical concept.

By Habermas's account, the capitalism of Marx's day was more or less the laissez-faire ideal that is spoken of today (and is being prescribed by Western industrial nations to the new economies of Eastern Europe). This mode of production was legitimized by the ideology of 'just exchange': the picture of the market that claimed that it was essentially a center of 'pure procedural justice'; that the free market itself was non-judgemental and morally neutral; and that any inequalities that resulted from the capitalist mode of production were a function not of the market itself, but of the industriousness and ingenuity (or lack thereof) of those who participated in the exchange of goods or labour for capital. The core of Marx's critique of bourgeois ideology was the claim that this picture of exchange on the free market as being a purely quantitative manner masked a qualitative difference between the position of the worker selling her or his labour and the capitalist buying it in exchange for capital. It was in this qualitative difference-- that between someone who must sell her or his labour to survive

and someone who may buy it in the interest of generating profit--the exploitative dimension of so-called 'just exchange' was hidden.¹⁵ Thus the ideology of just exchange worked to legitimize a mode of production, and the critique of ideology served to de-legitimize the same mode of production by showing that it represented what was in the interests of some (the maintenance of the free market) as being in the interests of all.

Habermas claims that as capitalism has evolved, so too has its mode of legitimation, and so too must change the critique of ideology. The most basic element of the shift in the capitalist mode of production lies in "the cumulative growth, on the part of the state, of interventionist activity which is designed to secure the stability and growth of the economic system."¹⁶ The recognition by government economic planners that capitalism, if left to run its course, would produce the type of radical disparity of income that Marx predicted led to state intervention into the economy in the interests of preserving the relative stability of the economic system¹⁷. Habermas argues that

¹⁵Cf. Karl Marx, "The Fetishism of Commodities and the Secret Thereof", Capital (Chicago: Charles Kerr & Co., 1906), pp. 81-96.

¹⁶Habermas, "Some Difficulties in the Attempt to Link Theory and Practice, Theory and Practice (Boston: Beacon Press, 1973), pp. 3-4.

¹⁷The problem with this eminent radical disparity of income was not a moral one, but a practical one. An extremely impoverished working class that increased in size as the owners of the means of production became wealthier and fewer in number would result in a state of underconsumption: there would not longer be enough people with enough money to purchase the goods being produced. State intervention in the economy was not

applying this solution (i.e., state intervention into the economy) to the impending crisis of capitalism has led to the development of the public realm into a space in which the relevant questions raised and solutions offered are of an exclusively technical nature. Capitalism no longer appears as a system in need of legitimation, but, rather, a process in need of technical repair. The consequence of this shift, Habermas argues, has been the exclusion of practical¹⁸ matters from the public realm and the subsequent depoliticization of the public. So ideology is no longer called upon to provide practical legitimation of the capitalist mode of production (e.g., to represent the free market as the seat of 'just exchange'), but, rather, ideology acts to de-politicize the public by displacing practical questions from the public realm through demanding the relevance and primacy of technical rather than practical issues.

Thus Habermas is led to claim:

The immunizing power of ideologies, which stifle the demand for justification raised by discursive examination, goes back to blockages in communication, independently of the changing semantic contents. These blocks have their origin within the structures of communication themselves.¹⁹

motivated by benevolence, but by interest in self-preservation.

¹⁸Here, and in all his writings, Habermas is using 'practical' as Kant did (and Aristotle before him), not as a synonym for 'prudential.' Thus, in the context above, the 'technical/practical' is analogous to the 'prudential/moral' distinction more common in Anglo-American moral theory.

¹⁹Habermas, "Some Problems in the Attempt to Link Theory and Practice," p. 12.

Let me unpack this. The power of ideology (i.e., in its contemporary form) is independent of semantic content because it no longer seeks to legitimize the extant mode of production by providing an erroneous justification, but, rather, acts to exclude questions of justification from the public realm. In this sense, the blockages in communication "which stifle the demand for justification" are found in the very structure of communication; that is, in a public language game structured to focus on technical problems at the exclusion of practical issues. Thus, 'ideology' appears in Habermas as 'systematically distorted communication.'

Ideology in this sense (rather than in the classical sense, as explained in (1), above) functions as the pathological because it effectively removes from public space those concerns which admit of universalization. Technical questions, by their very nature, take as their answers hypothetical imperatives. Insofar as this modern ideology acts to stifle the raising of questions concerning universalizable interests, it assumes, for Habermas, a position analogous to that which Kant gave material interests.

Now, as the "veil of ignorance" served as an antidote for the intrusion of *contingent interests* in Rawlsian meta-ethics, the "ideal speech situation" serves as something of an antidote for *ideology* in Habermas's meta-ethics. To understand this we have to explore the implications of Habermas's analysis of ideology.

It follows from Habermas's analysis of the form ideology has

taken in advanced capitalism that 'ideology' cannot be understood simply in contrast to 'truth.' Ideology, roughly speaking, is not a matter of the misrepresentation of reality, but rather the deformation of the process by which we come to accept descriptions of the world or prescriptions for our actions. Rather than standing in opposition to *true description*, ideology, in Habermas's account, is opposed to *just procedure*.

An obvious and strong parallel with Rawls emerges here. Rawls argued that once we removed certain barriers by imposing certain constraints--i.e, depriving agents of knowledge of the morally arbitrary--, the process of consensus-building becomes rational. The criterion by which a principle is judged as binding (or not) is whether it can be the object of rational consensus (whether it can be, that is, universalized).²⁰ What stands in the road of justice, for Rawls as well as for Habermas, is a *procedural* pathology. This is the link between the two revisions of Kantian meta-ethics that I suggested Rawls and Habermas shared: a reinterpretation of Kant's ethics in procedural terms, and a revision of Kant's conception of the pathological that accepted Kant's identification of pathology with contingency, but rejected his identification of the contingent with material interest.

But if we consider more carefully what constitutes procedural pathology for Habermas, two important differences

²⁰rationality = impartiality, impartiality = fairness, fairness = justice.

appear between Habermas and Rawls.

(1) Habermas's meta-ethics is more purely procedural than Rawls's. Consensus is deviant for Rawls when it is informed by particular facts. The veil of ignorance acts to bar the entry of these facts in the consensus-building process. By contrast, consensus is deviant for Habermas when the process of consensus-building *itself* is constrained. As Thomas McCarthy says on behalf of Habermas,

The guiding light is that a consensus is "rationally motivated" or "grounded" if brought about solely through the cogency of the arguments employed (and not, say, through external constraints on discourse or thorough "internal" constraints built into the structure of discourse).²¹

The "ideal speech situation" serves not to exclude contingent or morally arbitrary knowledge (as the veil of ignorance does), but to free the process of consensus-building from structural constraints²². McCarthy continues:

[T]he conditions under which rational consensus is possible--what Habermas calls the "ideal speech situation"--must ensure not only unlimited discussion but discussion that is free from distorting influences, whether their source be open domination, conscious strategic behaviour, or the more subtle barriers to communication deriving from self-deception.²³

²¹Thomas McCarthy, The Critical Theory of Jürgen Habermas (Cambridge: MIT Press, 1978), pp. 304-305.

²²The sense in which the ideal speech situation "serves" to free the consensus-building process from structural constraints will be explained below.

²³McCarthy, p. 306.

Habermas's meta-ethics is thus, in this way, more purely procedural than Rawls's because for Habermas 'rational consensus' is characterised purely in terms of the procedure by which it is attained, rather than by the information which is available to the agents wishing to reach agreement.

(2) Procedural pathology has a deeper significance for Habermas than for Rawls because for Habermas being the object of rational consensus is not only the criterion of the just but also the criterion of the true. Habermas's theory of truth draws more directly on Peirce than on Kant.²⁴ Peirce's clearest statement of his theory of truth emerges in his discussion of the nature of reality in "How to Make Our Ideas Clear": "The opinion which is fated to be ultimately agreed to by all who investigate, is what we mean by the truth."²⁵ (Peirce continued: "...and the object represented in this opinion is the real.") Habermas's version is:

I may ascribe a predicate to an object if and only if every other person who *could* enter into a dialogue with me *would* ascribe the same predicate to the same object. In order to distinguish true from false statements, I make reference to the judgement of others--in fact to the judgement of all others with whom I could ever hold a dialogue (among whom I counterfactually include all

²⁴though Habermas does not state any particular fidelity to Peirce. Apel, another contemporary representative of the Frankfurt School is much more explicit concerning the debt this conception of truth owes to Peirce.

²⁵Charles Sanders Peirce, "How To Make Our Ideas Clear", Justus Buchler (ed.), Philosophical Writings of Peirce (New York: Dover Publications, Inc., 1955), p. 38.

Peirce's theory of truth and its relationship to Kant will be discussed in chapter three, when I turn to examine Apel, who draws heavily on Peirce.

the dialogue partners I could find if my life history were coextensive with the history of mankind). The condition of the truth of statements is the potential agreement of all others.²⁶

This a form of what has come to be known as 'ideal limit theory.' Habermas's theory truth can be characterized as a *procedural* interpretation of ideal limit theory. By this I mean to distinguish between, on the one hand, an interpretation of the ideal limit that defines truth in terms of, say, the content of total science at the limit of inquiry, and, on the other hand, an interpretation of the ideal limit that defines truth in terms of the conditions that would obtain in an ideal situation of inquiry. If we are to affix the ideal limit theory tag to Habermas²⁷, it must be clear that 'limit' in 'ideal limit' does not refer to the limit of inquiry in the sense of the content of the most complete description of the domains of our investigation. Rather, 'limit' refers to the idealized conditions of perfectly free inquiry. The ideal limit here refers not to the book written at the end of inquiry, but the conditions that must obtain for the predicates 'true,' 'good,' 'just,' etc. to be conferred on the contents of the book.

²⁶Habermas, "Wahrheitstheorien," in Wirklichkeit und Reflexion: Festschrift für Walter Schulz (Pfullingen, 1973), p. 219. Translated by and cited by McCarthy, p. 299.

²⁷I put the point in these terms because Habermas does not use the ideal limit theory terminology. I use the term here for convenience, because, given the interpretation of the 'ideal limit' I give below, I think it captures much of Habermas's account of truth.

As a procedural interpretation of ideal limit theory, Habermas's theory of truth must have a normative understanding of 'consensus.' In the passage quoted from "Wahrheitstheorien" above, Habermas would perhaps have given a more complete description of his theory of truth had the last sentence read: "The condition of the truth of statements is the potential *rational* agreement of all others." The procedural interpretation of ideal limit theory (hereafter ILT_p) is coextensive, in terms of normative criteria, with a rational consensus theory of truth.

The 'ideal speech situation' emerges out of Habermas's theory of truth. Before explaining the link between the ideal speech situation and Habermas's neo-Peircean theory of truth, however, let me close the point that led us into Habermas's form of ideal limit theory: how Habermas's employment of the criterion of 'being the object of rational consensus' to truth as well as justice marks an important difference between Habermas and Rawls (outside of the obvious point that Rawls does not extend the 'rational consensus' criterion beyond ethics into epistemology). Habermas's theory of truth is grounded in two claims:

i. "Our first sentence expresses unequivocally the intention of universal and unconstrained consensus."²⁸ Habermas argues that the act of speaking involves the speaker in implicitly three validity claims:

the rightness that the speaker claims for his action in

²⁸Habermas, "Knowledge and Human Interests: A General Perspective", Knowledge and Human Interests (Boston: Beacon Press, 1971), p. 314.

relation to a normative context (or, indirectly, for these norms themselves); the truth-fulness that the speaker claims for the expression of subjective experiences to which he has privileged access; finally, the truth that the speaker, with his utterance, claims for a statement (or for the existential presuppositions of a nominalized proposition).²⁹

All these claims to validity admit of empirical testing. The only criteria available for such testing are those that we devise. These criteria, as well as both the particular tests in which they are employed and investigators which use them, are fallible. The only tribunal against which the adequacy of these tests may be judged is the tribunal formed by other--in fact all other--inquirers. The most positive result of such a test is consensus. Therefore, the very act of speaking presupposes, and takes as a necessary precondition, the desire to reach consensus.

ii. It follows from this that "the truth of statements is based on the realization of the good life."³⁰ What Habermas means here is that if the criterion of a proposition's being true is that the proposition would be the object of rational consensus (and we have already seen that the desire to reach consensus is bound up with the act of speaking), then the criterion of truth has normative consequences beyond epistemology. 'Rational consensus' has both epistemic and practical components. As we saw above, Habermas argues that a given consensus is rational if

²⁹Habermas, The Theory of Communicative Action, Vol.1 (Boston: Beacon Press, 1981), p. 307.

³⁰Habermas, "Knowledge and Human Interests: A General Overview", p. 314.

and only if it is brought about by the cogency of arguments and does not occur as a product of (other) constraints on argumentation. The structuring of arguments and the criteria of valid argumentation are epistemic matters; but ensuring that agents engaging in discussion enjoy freedom from coercion is primarily a *political* matter. Rational consensus is fully possible only within a society characterized by equality and freedom from domination and coercion. In this sense the truth of sentences is based on the anticipation of the realization of the good life.

So, as I claimed above, procedural pathology has a deeper significance for Habermas than for Rawls because for Habermas being the object of rational consensus is not only the criterion of the just but also the criterion of the true. This deeper significance lies not only in the fact that Habermas applies the criterion of 'being the object of rational consensus' more broadly than does Rawls, but, more importantly, because the application of this criterion epistemically has, in Habermas's account, normative consequences that extend beyond epistemology into ethics and political philosophy.

An argument for the ideal speech situation may be generated from the material covered above.³¹

P₁: Any particular application of any particular evaluative criterion is fallible.

P₂: The only available test we have for the cogency of the

³¹In my reconstruction of Habermas's argument, I am drawing loosely on McCarthy, pp. 307-308.

evaluation of propositions with respect to their validity is the scrutiny of the relevant speech community.

Therefore, truth claims have as their final tribunal the potential agreement of all members of the relevant speech community.

P₃: If agreement reached in critical discussion is to provide a warrant for truth claims, there must be a way to distinguish a rational consensus from a merely de facto consensus.

P₄: We have available to us such a criterion of demarcation: A particular consensus is rational iff it is brought about solely through the arguments employed and not through forces of coercion³² that are either internal or external to the discussion.

Therefore, speaking of propositions as being true commits us to positing an idealized state of inquiry, where participants in a given inquiry defend and criticize claims in a debate free from structural constraints: an ideal speech situation.

We might add, as well:

P₅: The act of speaking involves implicitly raising several validity claims whose truth value admits of empirical testing.

Therefore, engaging in speech commits us to the idea of an ideal speech situation.

The weakest point in this argument is P₂. I have not offered any arguments against discourse- or community-independent criteria for truth. Neither, as far as I know, has Habermas.³³ This is the weak point in ILT_p. Philosophers of this

³²By "internal coercive forces" I mean what Habermas takes contemporary ideology to be: a systematic exclusion of certain types of questions from what is taken to be legitimate public debate. "External coercive forces" are more explicit: laws barring types of speech, curfews, physical force, etc.

³³Though he has recognized the problem; cf. "Some Difficulties in the Attempt to Link Theory and Practice", p. 19. McCarthy reviews some of Habermas's responses to criticisms of consensus theories of truth, pp. 303-304, but does not offer on Habermas's behalf a positive argument for consensus rather than, say, correspondence theories of truth. Habermas may well have

persuasion--including Apel and Putnam along with Habermas--all ask us, in one way or another, to either abandon truth for rational warrant, or to equate truth with rational warrant. The weaker version of this appeal is to claim that a necessary condition for a proposition being true is that it is the object of rational consensus, and that if a proposition enjoys such vindications, we are warranted in holding it to be true. The stronger version is that this criterion offers a sufficient as well as a necessary condition for truth.³⁴ ILT_p, as a purely epistemological theory of truth, needs the stronger version, because it is only with such a claim that the claims of metaphysical theories of truth can shown to be irrelevant, if not incorrect.³⁵

such an argument; I have read only a fraction of his works, and only a fraction of his writing has been translated into English. However, none of the major critical literature on Habermas offers such a positive argument on Habermas's behalf.

³⁴Putnam gives his explicit endorsement to this stronger thesis: "[T]he notion of a fact (or a true statement) is an idealization of the notion of a statement that is rational to believe."

Hilary Putnam, Reason, Truth, and History (Cambridge: Cambridge University Press, 1981), p. 201.

I will give Putnam more detailed attention in my discussion of Apel in chapter three.

³⁵If only the weaker version were being defended, then ILT would be completely compatible with a metaphysical realist theory of truth. That is, if 'being the object of rational consensus' were only a necessary, but not sufficient, criterion for truth, the metaphysical realism would have space to add, as a second necessary condition, that the proposition accurately describe the mind- and discourse- independent world. If such a correspondence picture were accurate, and if the means to make such descriptions were available, then it would seem this criterion would override the rational consensus criterion.

The final upshot of Habermas's revision of Kant--as it was with Rawls's--is a naturalization of Kant's account of autonomy. Noumenal autonomy is replaced with the autonomy enjoyed by persons in a situation of speech free from coercion. 'Autonomy' undergoes a strong reinterpretation here, insofar as the test of the validity of a normative claim can no longer be contained within one person's reflection, but must take as its tribunal the rational consensus of all parties which it may effect. McCarthy explain this shift clearly:

Habermas's discourse model represents a procedural reinterpretation of Kant's categorical imperative: rather than ascribing as valid to all others any maxim that I can will to be a universal law, I must submit my maxim to all others for the purposes of discursively testing its claim to universality. The emphasis shifts from what each can will without contradiction to be a general law, to what all can will in agreement to be a universal norm.³⁶

This shift does not thereby mark the subordination of autonomy to intersubjectivity, but a reinterpretation of the conditions of autonomy. Autonomy in Kant was most fundamentally autonomy from material interests, which, for Kant, represented the pathological determination of the will because he felt they did not admit of generalization. However, as I mentioned above, for Habermas an interest that proves to be merely individual is rendered unsuitable as a basis for universal legislation not qua interest, but qua non-generalizable. In McCarthy's words, autonomy "is no

³⁶McCarthy, p. 326.

longer defined by way of opposition to interest per se"³⁷, but rather in opposition to those interests which cannot be generalized. For Habermas, autonomy is autonomy from the discursive and political structures wherein the interests of some are presented and enforced as the interests of all.

I will turn now to review and conclude my presentation of Rawls and Habermas by way of an assessment of the boundaries of the Kantian interpretation I have given them, and an examination of the answers their meta-ethics provide for the two questions I asked of neo-Kantianism at the end of the last chapter.

Rawls and Habermas evaluated

(1) How far may I push my Kantian interpretation of Rawls and Habermas? That is, have I overstated their debt to Kant (and, as such, their common ground)?

I have argued that at the heart of both Rawlsian and Habermasian meta-ethics are two revisions of Kant: a naturalizing of Kant's account of autonomy, and a reinterpretation of Kant's account of the pathological. In both cases and with both writers, these revisions are substantial. The question is: Are these revisions so substantial that they are better understood as points of departure from the Kantian tradition rather than, as I have argued, as points of continuity?

³⁷McCarthy, p. 328.

Oliver A. Johnson has argued as much with respect to the Kantian interpretation Rawls gave to the meta-ethics behind his theory of justice. Johnson argues that Rawls's interpretation of autonomy, the categorical imperative, and rationality, all suffer from a misconstrual of Kant that is serious enough to draw into question the appropriateness of characterizing Rawlsian meta-ethics as Kantian. Since my focus here has been on autonomy, and--as Steven Darwall pointed out in a response to Johnson³⁸--Johnson's arguments against Rawls's account of the categorical imperative and rationality derive from Rawls's account of autonomy, I will concentrate on what Johnson takes to be the shortcomings of a Kantian interpretation of autonomy as secured behind the veil of ignorance.

Johnson correctly notes that "[t]he inference Rawls wants us to make is that, since decisions motivated by specific, contingent wants constitute heteronomy, decisions in which these can play no part must be autonomous."³⁹ The problem, Johnson argues, is that the motivations of persons in the original positions do not change behind the veil of ignorance; rather, "only the conditions under which they reach their decision has been changed."

They still choose the principles they do motivated by the desire to promote their own interests in the best way they can. Each man's decision, in other words,

³⁸Steven Darwall, "A Defence of the Kantian Interpretation", Ethics 86:2 (1976), pp. 164-170.

³⁹Oliver A. Johnson, "The Kantian Interpretation", Ethics 85 (1974), p. 61.

still has its basis in want and desire.

For Kant the distinction between autonomous and heteronomous acts lies not in the circumstances in which the acts are performed, but in the motive from which they are performed. An action originally heteronomous is not rendered autonomous, even though performed under a veil of ignorance, if the nature of its motivation is unchanged.⁴⁰

So Johnson's point is that Rawls has misconstrued Kant's account of autonomy by understanding autonomy as opposed to contingent interest, rather than opposed to desire in general. This misconstrual, by Johnson's estimation, undermines Rawls's claim for a place for his theory in the Kantian tradition.

This objection would hold against Habermas, as well. As McCarthy claims on behalf of Habermas:

The aim of discourse is to come to a consensus about which interests are generalizable. In this construction, individual wants, needs, desires, and the interests need not--and cannot--be excluded, for it is precisely concerning them that agreement is sought.⁴¹

As I explained above, for Habermas, particular interests are unsuitable for universal legislation not because they are interests, but because they are non-generalizable. So Johnson would argue against giving Habermas's interpretation of autonomy a Kantian reading as well.

⁴⁰Johnson, p. 62. Johnson's points against Rawls's reinterpretation of the categorical imperative and Kant's theory of rationality follow from this: if choices under the veil of ignorance are heteronomous, then they can yield only hypothetical imperatives; if choices under the veil of ignorance are motivated by desire, they are not purely rational.

⁴¹McCarthy, p. 327.

Against Johnson, I am claiming that Kant's equation of heteronomous or pathological determination of the will with interested determination of the will is secondary to Kant's equation of the pathological with the contingent. I am suggesting, that is, that Kant was concerned first of all to exclude the contingent from the moral, and *then*, upon deciding that the contingent lay in material interests, equated the pathological or heteronomous with the sensuous or the interested. It is on this second point that Rawls and Habermas depart from Kant, but on the first--the equation of the pathological with the contingent--they are in agreement.

There is textual backing for my position. As Darwall points out in his reply to Johnson⁴², in the Groundwork Kant claims:

I understand by 'kingdom' a systematic union of different rational beings under common laws. Now since laws determine ends as regards their universal validity, we shall be able--if we subtract from the personal differences between rational beings, and also from all the content of their private ends--to conceive of a whole of ends in systematic conjunction.
(Groundwork, p. 101)

For Kant, these "personal differences" and the "contents of private ends" were to be found in material desire. For Rawls, they are to be found in interests tied to knowledge of your position of society; for Habermas, in individual interests that do not admit of universalization. The primary concern is with the exclusion of the contingent, whatever that may turn out to

⁴²I am not siding with Darwall's response to Johnson, but crediting him for pointing out this passage.

be. I am arguing that what follows from this is that Rawls's and Habermas's re-interpretations of Kantian autonomy maintain fidelity to the most fundamental elements of Kant's meta-ethics.

The question of the relative fidelity to Kant serves to highlight an important difference between Rawls and Habermas. Rawls is 'more Kantian' than Habermas insofar as Rawls, with the veil of ignorance, seeks to exclude certain facts from the legitimate content of moral deliberation. Kant sought to exclude facts that pertained to the achievement of material interests. Habermas is 'more Kantian' than is Rawls insofar as Habermas, like Kant, offers a common ground for the vindication of both epistemic (theoretical) and moral (practical) claims: 'rational consensus' or "the agreement of free citizens." In this sense, Habermas is also more Kantian than Rawls insofar as we find in Habermas a contemporary version of the epistemic/political interface that I argued was to be found in the critical philosophy.

(2a) Can we reformulate Kant's account of the pathological and keep his account of the normative?

Normativity arises in Kant's meta-ethics out of the tension between the kingdoms of the earth and the kingdom of ends, or between our residence in the phenomenal realm and our residence in the noumenal realm. This account of normativity is intimately tied up with Kant's account of the pathological. Morality in Kant was possible only through freedom from pathological or heteronomous determination of the will; securing this freedom

required positing the noumenal realm. The question that arises is: can we reformulate the pathological in such a way that freedom from the pathological determination of the will does not require positing the noumenal realm but nevertheless avoid losing the grounding of normativity with the loss of the noumenal/phenomenal distinction?

In a word: yes. As we saw, both Rawls's and Habermas's reformulations of the pathological led to the formulation of idealized states of autonomy. Freedom from knowledge of the morally contingent gives us the original position; freedom from the distorting effects of public language games that distort our means of generalizing interests gives us the ideal speech situation. Both idealized states of autonomy function as regulative ideals directing moral deliberation and grounding political critique. The normative in Rawls and Habermas arises as it does in Kant: from the tension between the imperfect state of autonomy in which we live, and the idealized state of autonomy toward which our practices point.

(2b) Can naturalizing Kantian autonomy, while escaping the positing of a noumenal realm, simultaneously avoid taking recourse to a transcendental vindication of the authority of reason?

Let me explain how and why this question arises. By naturalizing Kantian autonomy, Rawls and Habermas make room for ethics within what Kant called the phenomenal realm. I have argued above that Kant's account of normativity can be recast by

analogy along with this naturalization: crudely, by replacing the noumenal realm with the original position or the ideal speech situation. Recall, however, as I argued toward the end of the last chapter, that Kant's vindication of the authority of reason lay somewhere between the transcendental and the political: while, on the one hand, the "agreement of free citizens" was possible only in a society characterized by tolerance, on the other, this 'freedom' was not an illusion only because we were transcendently free. So the question for Rawls and Habermas is: Can an analogy for transcendental freedom be constructed in the same manner as the analogies for the noumenal realm that, like the veil of ignorance and the ideal speech situation, do not push us into the transcendental or a priori?

This time, the answer is no. Let me place the question in more contemporary terms. Transcendental freedom, for Kant, guarantees that all persons are ultimately responsible for their actions. The existence of transcendental freedom removes the threat that practical freedom is an illusion and vindicates our practice of imputing responsibility to the actions of persons. That we are transcendently free makes it impossible for us to 'opt out' of taking responsibility for our actions and to avoid judging those actions by the rational dictates of the categorical imperative. Do Rawls's and Habermas's meta-ethics also contain such a prohibition on escape clauses?

Stephen K. White phrases this question well by noting that Habermas's 'communicative ethics' (what I referred to above as

the extra-epistemic normative implications of the rational consensus theory of truth) must be binding on first-person dictators and systematic free-riders.⁴³ Recall that above I mentioned that Habermas argued that the act of speaking involved implicitly making several claims to validity; the question here is how is it that I am bound to defend these claims. The challenge lies in the claim that

[i]f there is any obligation flowing from engaging in communicative action, then it is an obligation which one *chooses* to take on; one could just as easily choose to avoid communicative action altogether in the future and orient oneself exclusively toward strategic action, thereby avoiding normative obligation.⁴⁴

Habermas would respond to this challenge by claiming, on the one hand, that the obligations implicit in communicative action are transcendental conditions of such action, and, on the other, that "other forms of social action--for example, conflict, competition, strategic action in general--are derivatives of action oriented to reaching understanding [i.e., communicative action]."⁴⁵ The transcendental status of the obligations raised by speaking is marked by the fact that an actor who refuses to acknowledge such constraints on communicative action is guilty of

⁴³Stephen K. White, The Recent Work of Jürgen Habermas (Cambridge: Cambridge University Press, 1988), p. 53.

⁴⁴White, p. 51. This passage contains terminology I have not introduced. 'Communicative action' is action aimed at understanding; 'strategic action' is action aimed at achieving some pre-determined goal.

⁴⁵Habermas, "What is Universal Pragmatics?", Communication and the Evolution of Society (Boston: Beacon Press, 1979), p. 1.

a "performative contradiction."⁴⁶ Habermas, unfortunately, is not very clear on the nature of this contradiction.⁴⁷ The point I set out to make here, however, has been made. Habermas must take recourse to the transcendental to prohibit the sceptic or nihilist from writing an escape clause into communicative ethics.

What about Rawls? Steven Darwall correctly suggests that without the Kantian interpretation Rawls can offer no compelling reasons for why anyone ought to be interested in justice. Darwall argues that Rawls's alternative justification, that the principles he offers would accord to our intuitions in 'reflective equilibrium', cannot accomplish this task.

[T]his justification of the principles [i.e., through 'reflective equilibrium] leaves unanswered the deeper question of why one should be interested in justice, even if it is true that our considered judgements about it can be organized by the principles. That is, it does not imbed a theory of justice in a theory of practical reason.

The Kantian interpretation suggests that there may be a deeper justification for the principles--namely, that they would be chosen from a perspective which, since it is that "procedural interpretations of Kant's conception of autonomy and the categorical imperative," it is compellingly rational to adopt.⁴⁸

⁴⁶Cf. White, p. 51. As we will see in the next chapter, the kind of analysis Habermas is suggesting points to what is taken as perhaps the most basic feature of the transcendental status of a principle or criterion: that it cannot be denied without being presupposed.

Habermas borrows the idea of performative contradiction from Apel, whose defence of transcendentalism I will be examining in the next chapter.

⁴⁷The idea seems straightforward enough, though. For example, a liar-type paradox occurs if you tell me that you refuse to acknowledge the truthfulness of your utterances.

⁴⁸Darwall, pp. 164-165.

But how far does this Kantian interpretation go? Even if one were to accept that rationality were a matter of impartial reasoning, and that if one reasoned impartially--behind the veil of ignorance--one would choose the principles of justice, there seems to be nothing preventing one from refusing to be rational, or, while recognizing the dictates of rationality, overriding them in the interests of greater personal power. In short, there is nothing in Rawls that prevents one from refusing to take responsibility for one's actions.⁴⁹ Rawls needs an analogue for Kant's transcendental freedom, some argument that proves it impossible to refuse the normative consequences of the ability to reason impartially.

Through my examination of Rawls and Habermas, I have tried to defend two claims: (1) that Kant's meta-ethics in general and his account of autonomy in particular have relevance for the concerns being addressed in contemporary political theory and are represented in the neo-Kantianism of Rawls and Habermas; and (2) that Kant's meta-ethics in general and account of autonomy in particular can be recast in light of contemporary concerns (and are so re-cast by Rawls and Habermas) and retain the formal

⁴⁹I mean 'taking responsibility' in the sense that Kant would have meant. Clearly, someone could take responsibility for overriding the principles of justice in the interests of greater power by simply declaring that to be the object of her or his actions. For Kant, however, such an agent would not be facing up to the responsibility demanded by the dictates of practical reason.

structure that I argued (in Chapter One) served as a grounding, for Kant, for radical political critique. We have seen, however, that the project of naturalizing Kantian autonomy--at least as it is represented in Rawls and Habermas--cannot be done without transcendental residue. That is, while you can replace the noumenal realm with the original position or the ideal speech situation, you cannot--without a transcendental freedom-analogue--prevent persons from recognizing the rationality of principles chosen either behind the veil of ignorance or the ideal speech situation and then choosing to 'opt out' of acting rationally. In Kant's terms, Rawls and Habermas cannot show why the Devil cannot be a logician.

It is to this concern that I turn in chapter three. I will explore the contemporary relevance of the transcendental first through an examination of the contemporary literature on transcendental arguments, and then through a defence of the procedural interpretation of ideal limit theory against Rorty's criticisms of neo-Kantian political philosophy.

CHAPTER THREE

AUTONOMY AND PROCEDURE

As promised in the introduction to this paper, I will now turn to consider various attacks made on the account of autonomy offered by contemporary neo-Kantian political philosophy. In this chapter I will address the criticisms that approach this account of autonomy from the transcendental side. In the fourth and final chapter, I will focus on critics of neo-Kantianism who approach the neo-Kantian account of autonomy from the political side.

In the last chapter I argued that, while Rawls's and Habermas's naturalization of Kantian autonomy successfully retained Kant's account of the normative and his grounding of radical political critique, this naturalized neo-Kantianism remained mute to the challenge of a particular kind of moral nihilist. Both Rawls and Habermas have at the core of their meta-ethics an inference from a *rational* 'ought' to a *moral* 'ought', and it is at the point of this inference that the moral nihilist raises her challenge.¹

As we saw, Rawls argues that the norms of impartial

¹Let me explain why I am affixing the 'nihilist' tag to the challenge to Rawls and Habermas. The position I have in mind, as I described in the previous chapter, and will review below, is one that refuses to acknowledge as morally binding the extra-epistemic normative consequences of a theory of rationality. I am calling this a type of moral nihilism because it involves denying at least a particular type of claim concerning the knowledge of moral truths or justification for moral rules.

reasoning give us the principles of justice². The constraints of the veil of ignorance guarantee that the consensus formed behind the veil is, in principle, universalizable, and thus morally binding. Habermas argues that "the truth of statements is based on the anticipation of the realization of the good life."³ According to Habermas, the criterion of the truth of a statement is that it be the object of rational consensus. 'Rational consensus', for Habermas, has extra-epistemic normative consequences that reach into ethics and political philosophy: a necessary condition of rational consensus is freedom from coercion. Hence, the norms of rationality imply the obligation to secure certain political conditions.

The most serious challenge to Rawls and Habermas, in this context, comes from anyone (whom I have identified with a particular form of moral nihilist) who, though accepting the norms of rationality that Rawls or Habermas offer, refuses to

²Let me emphasize, as I did in the last chapter, that my concern with Rawls is at the meta-ethical level. For the purposes of my argument, I am indifferent both to the content of Rawls's principles of justice and the question of whether or not they would, in fact, be favoured by impartial deliberators. Hare, in his review of TOJ, argued that the contractors behind the veil of ignorance would favour an alternative principle based on utility with a floor constraint. A forthcoming book by Norm Frohlich and Joe Oppenheimer claims to have empirically verified Hare's conjecture. Even if Hare, Frohlich and Oppenheimer were correct, my concerns would be unchanged. Rawls would have been incorrect about the content of the principles of justice, but his claim about the normative force granted to *whatever* would be chosen behind the veil of ignorance would be untouched. It is this latter meta-ethical claim that is the focus of my concerns.

³Habermas, "Knowledge and Human Interests: A General Perspective", p. 314.

admit the moral norms that Rawls and Habermas argue follow from their respective accounts of rationality. Rawlsian meta-ethics cannot prevent someone from refusing the politically normative consequences of the ability to reason impartially. Habermas offers a response to the analogous challenge to his own meta-ethics. To the nihilist who--as I described in the previous chapter--either (1) refuses the moral commitments implied by communicative action, or (2) refuses to engage in communicative action, Habermas responds by claiming (1) that these commitments are the necessary preconditions of communicative action, and (2) that all non-communicative forms of action are derivative of communicative action. Habermas denies that the escape route(s) proposed by the nihilist place that nihilist beyond the scope of the minimalist ethics implied by communicative action, on the grounds that the nihilist's challenge presupposes, in some way, what the nihilist is denying.⁴

Habermas's claim relies on what--as we will see below--many philosophers take to be a central feature of a transcendental argument: that what is proven by a transcendental argument cannot be challenged without being presupposed. Habermas's argument also points to the most basic difference between Kant's use of transcendental argumentation and the relevance of transcendental arguments to contemporary neo-Kantian political theory. For

⁴Recall from the last chapter Habermas's claim that the nihilist's denial results in a 'performative contradiction', a contradiction that I suggested was a form of the liar's paradox.

contemporary neo-Kantian political philosophers, the challenge to be defeated or diffused comes not from the epistemic sceptic, as it did for Kant, but from the moral nihilist. Kant sought to defeat or diffuse Hume's scepticism about causation by arguing that the assumption that every event was the effect of some cause is not a hypothesis, but rather a necessary precondition of possible experience. Kant's argument was transcendental because he claimed that the conceptual scheme comprised of the Categories --which included 'causality'--was not optional or dispensable, but represented the absolute preconditions of rationality⁵. By analogy, Habermas is forced to claim transcendental status for the preconditions of communicative action to prevent the moral nihilist from proposing and/or acting on an alternative (say, anarchistic) ethics⁶.

I think this recourse to the transcendental is the most significant problem for contemporary neo-Kantian political

⁵By no means should this be taken as an exhaustive description of a transcendental argument. The distinguishing features of a transcendental argument will be taken up in the first section of this chapter.

⁶Let me clear up any confusions that may arise from my suggestion that a moral nihilist may defend an ethics. As I described above, the position I have in mind is one that refuses to acknowledge as morally binding the extra-epistemic normative consequences of a theory of rationality, and is thus a form of moral nihilism insofar as this position involves denying at least a particular type of claim concerning the knowledge of moral truths or justification for moral rules. The alternative ethics such a nihilist might propose is not a comprehensive alternative, but something analogous to White's "first person dictator" or "systematic free-rider".

philosophy⁷. Recourse to the transcendental is a problem rather than a solution for neo-Kantians because--second only perhaps to the noumenal/phenomenal distinction--the linked notions of transcendental preconditions and transcendental arguments are the most disputed elements of the critical philosophy.

My argument in defence of the neo-Kantian account of autonomy proceeds regressively; that is, I start with a strong position--Kant's original formulation of transcendental arguments--and pare it down to a weaker position--Habermas's argument for the transcendental preconditions of communication. I will address the problems of the contemporary relevance of transcendental arguments in two stages, marked by the two sections of this chapter. In the first section, I will explain what I take to be the central features of a transcendental argument, drawing on both Kant and some of his recent commentators, and then turn to address what I take to be the most

⁷A bold statement, given that I am generalizing from two neo-Kantians, Rawls and Habermas. But I think this generalization is justified. Rawls and Habermas appear in the literature as spokespersons for neo-Kantian liberalism and the neo-Kantian left, respectively. Furthermore, I am including Apel and Putnam in this generalization. Most importantly, however, I take the kind of naturalizing that, I have argued, characterizes Rawls's and Habermas's programs to be the heart of contemporary neo-Kantian political theory (and in this, if nothing else, Rorty agrees with me), and I cannot see how this program can be completed without the kind of transcendental residue that, I have argued, is left over in Rawls's and Habermas's meta-ethics. If the type of moral nihilist with whom I am concerned can be defeated on naturalized grounds, all the better for friends of naturalism and those who resist the neo-Kantian tradition for its violation of the naturalist injunction against the transcendental and the a priori. However, I have not seen such a thorough naturalization of Kant, and doubt that one can be produced.

significant criticisms raised against transcendental arguments in the recent literature. Specifically, I will attempt to answer criticisms advanced by Barry Stroud and Richard Rorty's earlier, more technical objections to transcendental philosophy. I will argue that, while Stroud's and Rorty's criticisms may hold against certain neo-Kantian metaphysicians, the type of transcendental argument required by contemporary neo-Kantian political philosophy is, by and large, immune to the criticisms offered by Stroud and Rorty. In the course of my argument I offer a version of the type of neo-Kantian transcendental argument that, as I described above, seeks not to answer the doubts of the epistemic sceptic, but deny the efficacy of the refusals of a type of moral nihilist. The argument I will offer draws heavily on Habermas and the material I covered in the last chapter.

In the second section of this chapter I turn to the more general issues involved in transcendental political philosophy, focusing on Rorty's later, more polemical criticisms of neo-Kantianism. I accept some of Rorty's misgivings, and argue that his strongest points indicate that we ought to prefer the Habermasian version of naturalized Kantianism over Rawls's revisions. However, I argue that Rorty's criticisms do not hold against a generalized neo-Kantian position that I articulate drawing not only on Habermas but also Apel and Putnam. I argue that this generalized neo-Kantian position can be understood as a procedural interpretation of ideal limit theory. I argue,

finally, that the distance between Rorty's anti-transcendentalism and neo-Kantianism in the form of a procedural interpretation of ideal limit theory is not nearly as great as Rorty would suppose (and hope). That is, I argue that Rorty's characterization of the anti-transcendentalist pragmatism that he endorses is at least compatible with, and perhaps implies, the transcendentalism characteristic of ideal limit theory given the procedural interpretation of ideal limit theory that I argue is shared by Habermas, Apel, and Putnam.

I think that Rorty deserves such detailed attention for two reasons. (1) He appears in the literature as the most popular and articulate spokesperson for contemporary anti-transcendental philosophy. As such, his criticisms must be addressed in any defence of transcendentalism. (2) Rorty considers his position to be far removed from Kantian transcendental philosophy. I argue that, *pace* Rorty, his difference with neo-Kantianism comes to a fairly precise and technical point, a point that does not mark the gulf he envisions between his anti-transcendentalism and neo-Kantian political theory. As such, I offer my argument against Rorty as something of an appeal to the anti-transcendental currents of contemporary philosophy (especially naturalism and post-structuralism) that transcendentalism is both not so easily dismissed as it appears, and is not so far from the more methodologically and ontologically austere intuitions that are enjoying prominence in contemporary philosophy.

Transcendental arguments

This section is itself divided into three parts. In the first I describe what I take to be the essential characteristics of a transcendental argument (hereafter TA), drawing both on Kant and some of his contemporary commentators. The second part will be devoted to an evaluation of Stroud's criticisms of TAs. The final part will be given to a consideration of Rorty's misgivings about TAs.

What is a transcendental argument?

The first important point to make about transcendental arguments is that they are not distinguished by their form or logical structure⁸. As Stroud argues,

The term 'transcendental' is not to be thought of as naming another form to be put alongside *modus ponens*, *modus tollens*, and the others. Rather we understand what is special or unique about a transcendental argument when we appreciate the special status of the 'principles' that would be established by a successful transcendental proof.⁹

That is, the formal structure of a TA does not serve to distinguish it from other arguments because rather than defining or guaranteeing the transcendental status of the conclusion of a TA, this formal structure follows from the 'special status' of

⁸Pace Moltke S. Gram, "Must We Revisit Transcendental Arguments?", Philosophical Studies 31 (1977), pp. 235-248, and "Transcendental Arguments", Nous 5 (1971), pp. 15-26.

⁹Barry Stroud, "Transcendental Arguments and 'Epistemological Naturalism'", Philosophical Studies, p. 111.

transcendental principles.¹⁰ What, then, is the status of a transcendental principle?

In the First Critique, Kant declared,

I entitle *transcendental* all knowledge which is occupied not so much with objects as with the mode of our knowledge of objects in so far as this mode of knowledge is to be possible a *priori*. (A 11-12/B 25)

This can be understood only within the context of Kant's Copernican revolution. Impressed by the success of geometry since the Classical period and physics since Newton, Kant sought to extend the method employed by these sciences to the study of metaphysics. Concerning the then recent developments in physics, Kant ventured that

[physicists] learned that reason has insight only into that which it produces after a plan of its own, and that it must not allow itself to be kept, as it were, in nature's leading strings, but must itself show the way with principles of judgement based upon fixed laws, constraining nature to give answer to questions of reason's own determining. Accidental observations, made in obedience to no previously thought-out plan, can never be made to yield a necessary law, which alone reason is concerned to discover. (B xiii)

Kant argues that a similar revolution occurred earlier in mathematics, an intellectual revolution "far more important than the discovery of the passage round the celebrated Cape of Good Hope" (B xi). This revolution was occasioned when some mathematician, whose name has not been recorded (but Kant suggests it may have been Thales), attempted to articulate the properties of a triangle.

¹⁰As well, at least some element of the formal structure of a TA may be exhibited by arguments that do not have transcendental principles as their conclusions.

The true method, so he found, was not to inspect what he discerned in the figure, or in the bare concept of it, and from this, as it were, to read off the properties; but to bring out what was necessarily implied in the concepts that he had himself formed a *priori*, and had put into the figure in the construction by which he had presented it to himself. If he is to know anything with a *priori* certainty he must not ascribe to the figure anything save what necessarily follows from what he has himself set into it in accordance with his concept. (Bxii)

Kant suggests that we take the example set by the successes of mathematics and geometry and apply them to metaphysics.

Hitherto it has been assumed that all our knowledge must conform to objects. But all attempts to extend our knowledge of objects by establishing something in regard to them a *priori*, by means of concepts, have, on this assumption, ended in failure. We must therefore make trial whether we may not have more success in the tasks of metaphysics, if we suppose that objects must conform to our knowledge. (B xvi)

So, against the background of Kant's 'Copernican revolution' in metaphysics, we may return to the passage cited above concerning the nature of the transcendental¹¹. By calling transcendental "knowledge which is occupied not so much with objects as with the mode of our knowledge of objects in so far as this mode of knowledge is to be possible a *priori*", Kant meant to identify transcendental principles with those principles that explained how knowledge of objects was possible, given that empirical knowledge was not to be understood in terms of the assumption

¹¹Such a lengthy aside into Kant's Copernican revolution may seem unbalanced at this point, but the nature of this hypothesis and Kant's motivation for adopting it will be relevant below, when we examine the defence Kant may have had against Stroud's charge that TAs harbour covert verification principles.

that "all our knowledge must conform to objects," but rather, by assuming that empirical knowledge was possible only if "we suppose that objects must conform to our knowledge." In terms of a oft-cited metaphor, Kant agreed with the empiricists that the mind was a *tabula rasa*, but sought to explain how such a blank slate could be written upon, and argued that such an explanation must take recourse to *a priori* principles, because "[e]xperience teaches us that a thing is so and so, but not that it cannot be otherwise" (B 3). Or, in terms of a Quinean figure, Kant argued that experience marked the "boundary conditions" of knowledge¹², but that certain conditions must obtain for experience to be possible, that "pure *a priori* principles are indispensable for the possibility of experience" (B 5). These *a priori* principles--that is, those that explain the necessary conditions of possible experience--are transcendental principles.

TAs, then, in A.C. Genova's words,

in the *broadest* sense, demonstrate that a certain conceptual scheme having certain categorical features applies to and is logically presupposed for the possibility of a given domain of experience.¹³

If the given domain is what Kant called possible experience, then a TA will consist in showing that,

¹²Cf. Quine, "Two Dogmas of Experience", From a Logical Point of View (Cambridge: Harvard University Press, 1953), p. 42.

¹³A.C. Genova, "Good Transcendental Arguments", Kantstudien 75, p. 478. All further references to this paper will be made in the text with reference to the author and the page in parentheses.

with respect to the intelligible domain of possible experience, certain concepts are primitive, i.e., are non-trivially instantiated whenever any other concept is instantiated, or that certain propositions are core propositions, i.e., are true whenever any empirical propositions is true or false. (Genova, p. 480)

As Stroud suggests, TAs defend a position that is content neither with scepticism nor with conventionalism.

A sound transcendental argument therefore would show that it is wrong to think (with the conventionalist) that the only possible justification of our ways of thinking is "pragmatic" or practical, and equally wrong to think (with the sceptic) that they can be justified only by collecting direct empirical sense of their reliability.¹⁴

A sound TA would show that "it is wrong to think (with the conventionalist) that the only possible justification of our ways of thinking is 'pragmatic' or practical" because a sound TA would show that, while experience is possible only through the ordering of sensation under a conceptual scheme, there is a core set of concepts that enjoy a priori status. We may justify our ways of thinking by legitimating our employment of these concepts with a TA. A sound TA would show that it is wrong "to think (with the sceptic) that our ways of thinking can be justified only by collecting direct empirical sense of their reliability" because a sound TA would show that, while the boundaries of knowledge are coextensive with the boundaries of possible experience, experience itself is made possible only by its interpretation under a scheme of concepts. Appealing to experience to justify

¹⁴Stroud, "Transcendental Arguments", Journal of Philosophy 65 (1968), p. 244.

our way of thinking gets that problem backwards. Collecting empirical data cannot justify our way of thinking; rather, we can justify our ordering of experience by vindicating our use of a set of core concepts through a TA. Stroud's point can be rephrased in positive terms; Kant's epistemology lies somewhere between conventionalism and realism: while--to make allusion again to Quine--experience marks the boundary conditions of knowledge, such knowledge is made possible only through a conceptual scheme. There is, however, according to Kant, one unique legitimate conceptual scheme.

Thus TAs answer a question of justification; they serve to legitimate the employment of a set of core concepts. Kant tried to capture this justificatory function of TAs by naming the TA with which he defended the transcendental status of his Categories the 'Transcendental Deduction.' 'Deduction' did not mean for Kant what it does in contemporary logic. He explains his use of the term by way of analogy to jurisprudence.

Jurists, when speaking of rights and claims, distinguish in a legal action the question of right (*quid juris*) from the question of fact (*quid facti*); and they demand that both be proved. Proof of the former, which has to state the right or the legal claim, they entitle the *deduction*. (A 84/B 116)

Now among the manifold concepts which form the highly complicated web of human knowledge, there are some which are marked out for pure *a priori* employment, in complete independence of experience¹⁵; and the right to be so employed always demands a deduction. For since empirical proofs do not suffice to justify this kind of employment, we are faced by the problem how these

¹⁵This claim was already justified in the Preface: the Copernican revolution.

concepts can relate to objects which they do not obtain from any experience. The explanation of the manner in which concepts can thus relate *a priori* to objects I entitle their transcendental deduction.¹⁶
(A 85/B 117)

In light of Kant's explanation of his use of the term 'deduction', Jay Rosenberg is led to suggest that, insofar as the conclusion of a transcendental deduction states, "not a matter of fact, but a matter of right," it is

in other words, to say that something may be done. It articulates a principle of permission. The conclusion of a transcendental argument is thus to be a normative conclusion.¹⁷

The sense in which TAs serve as refutations of epistemic scepticism follows from the paradoxical position in which they place those who deny their conclusions. In Stroud's words,

According to [the sceptic], any justification for our beliefs [concerning the existence of the material world] will have to come from within experience, and so no adequate justification can be given. Transcendental arguments are supposed to demonstrate the impossibility or illegitimacy of this sceptical challenge by proving that certain concepts are necessary for thought or experience.¹⁸

Demonstrating that certain concepts are necessary for experience removes these concepts from doubt that is based upon the underdetermination of these concepts by experience. In this

¹⁶This lends support to my earlier claim that TAs cannot be specified in formal terms. Kant's use of the term 'deduction' specifies a type of argument in terms of its content or purpose, rather than in terms of its formal structure.

¹⁷Jay F. Rosenberg, "Transcendental Arguments Revisited", Journal of Philosophy 72 (1975), p. 612.

¹⁸Stroud, "Transcendental Arguments", p. 242.

sense Kant thought he had defused Hume's scepticism about causality because Kant thought Hume's misgivings about causality were misplaced. Hume argued that nothing in experience could assure us that a particular effect followed from some cause. Kant replied by claiming that one of the very conditions of experience was that objects appeared to us in the relationship of cause and effect. Hume, by Kant's account, got the problem backwards.

Following this line of argument, P.F. Strawson claims that

He [the sceptic] pretends to accept a conceptual scheme, but at the same time quietly rejects one of the conditions of its employment. Thus his doubts are unreal, not simply because they are logically irresolvable doubts, but because they amount to the rejection of the whole conceptual scheme within which alone such doubts make sense.¹⁹

In the context of Kant's response to Hume, we may say, in Strawson's words, that Hume rejected the conceptual scheme within which his doubts made sense insofar as Hume's scepticism about causality was grounded in the observation that 'causality' was underdetermined by experience, while 'causality' is, in fact, one of the concepts that must be employed to make experience possible. Hume thus suspended this conceptual scheme to doubt an element of the experience that that conceptual scheme makes possible.

¹⁹.P.F. Strawson, Individuals (London, 1959), p. 35.

It should be noted, however, that, as Stroud points out²⁰, the sceptic's questions are not answered or satisfied by such a reductio refutation. We cannot really say that the sceptic is refuted by such arguments, but rather that her position is defused. That is, a TA does not answer the sceptic by proving the existence of that about whose existence the sceptic raises her doubts. Rather, a TA defuses the sceptic's challenge by arguing that the challenge itself (or its formulation) is flawed.

In summary, we can say that a TA is an argument that purports to demonstrate that a certain categorical scheme is a necessary precondition of some domain of experience, and, by doing so, answers the *quid juris*, i.e., legitimizes the employment of this (particular) set of concepts. It follows from a successful TA that certain types of sceptical claims are defused. A TA shows that doubts raised about the legitimacy of a particular concept--a concept the sceptic regards as a *hypothesis*, as Hume regarded causality--that are based on the underdetermination of the concept of empirical data are misplaced, insofar as the concept (among others) serves as a precondition of some domain of experience. Appealing to the underdetermination of a (so-called) hypothesis by experience poses the questions backwards.

Kant's TAs purported to show even more than this however. He claimed that "the conditions of the *possibility of experience* in

²⁰Stroud, "The Significance of Scepticism", p. 294.

general are likewise the conditions of the *possibility of the objects of experience*" (A 158/B 197). This is what Genova calls the 'objectivity thesis': the claim that "there cannot be anything satisfying our requirements for experience without it also being the case that at least part of this is experience of objects and events which exist independently of the perceiver and conform to the specifications of the unique [conceptual framework]" (Genova, p. 480). In short, a full-fledged TA will have to close the gap between the necessary subjective conditions of experience and the nature of that which is experienced, or the sceptic will have logical space in which to re-present her doubts. Stroud's argument that TAs harbour covert verification principles is based on the claim that this gap cannot be closed.

The 'covert verification principle' criticism

A.C. Genova argues, correctly I think, that a TA is comprised of three steps. The first, which Genova calls a hypothetical deduction (HD), is "directed to justifying a particular conceptual framework...as a necessary presupposition of some contingent domain of possible experience" (Genova, p. 479). This part of the argument takes the form of a conditional proof. However significant such an argument is, it has limitations. As Stroud points out,

It is not a sufficient refutation of the sceptic who doubts that *p* to present him only with a conditional to the effect that if *not-p* we couldn't possibly do *A*. What is in question is whether we ever "validly" or

"justifiable" do A.²¹

In this sense, Kant's *quid juris* must not only secure the right to employ certain concepts, but must argue that the (epistemic or moral) activity grounded in these concepts is legitimized by virtue of its grounding in these concepts. This latter step is what Genova--borrowing from Kant--calls a metaphysical deduction (MD), "which provides a priori justification of a unique [conceptual framework]--a [conceptual framework] which is a necessary presupposition of all possible contingent interpretations of experience" (Genova, pp. 479-480). The final step is the most controversial. This step of the argument--the transcendental deduction (TD) proper--must establish what I referred to above as the 'objectivity thesis': in Kant's words, a TD must show how "*subjective conditions of thought can have objective validity*" (A 89/B 122).

It is this final step--from a MD to a TD--that Stroud argues is possible only with a covert verification principle. Rorty captures the first part of Stroud's argument succinctly in a reply to Stroud.

For the purposes of giving "X" [an object in the domain of experience whose reality we are justified in claiming on the basis of the vindication given this judgement by a TA] a place in the language-game, so to speak, apparent X's are as good as real X's. So, if all the "transcendentalist" has to go on is that "X" does have such a place, he will never get from there to

²¹Stroud, "Transcendental Arguments", p. 244.

the reality of X's.²²

It follows from this, Stroud argues, that

[r]ather than dealing with the conditions of *knowledge*, then, those conditionals [that make up the first steps of a TA] must assert that the truth of what the sceptic doubts or denies is a necessary condition of the *meaningfulness* of the doubt or denial.²³

Stroud continues,

Any opposition to scepticism on this point would have to rely on the principle that it is not possible for anything to make sense unless it is possible for us to establish whether S is true...hence the meaning of a statement would have to be determined by what we can know.²⁴

This is a version of the verification principle. In Kantian terms, Stroud's point is that we cannot get from the subjective necessity to the objective validity of a conceptual scheme unless we make some essential connection between our use of concepts (meaning) and the conditions whereby those concepts obtain in the world (truth). A TA, Stroud concludes, must have a covert verification principle that links meaningfulness and truth in a way that securing the conditions of the former means securing the conditions of the latter. If TAs harbour covert verification principles, then there are two problems: (1) Many years of tinkering by the Vienna School and its followers did not yield a version of the verification principle that could sustain itself

²²Richard Rorty, "Verificationism and Transcendental Arguments", *Nous* 5 (1971), pp. 4-5.

²³Stroud, "Transcendental Arguments", p. 251.

²⁴Stroud, "Transcendental Arguments", p. 255.

and do justice to the practice of science; and (2) if TAs harbour covert verification principles, then--according to Stroud--the TA's that house them become superfluous.

I do not plan to deal with these two latter claims here, though I believe that the former is telling and that the latter is sound. Nor will I examine whether Stroud's criticism holds against the neo-Kantian metaphysicians (Strawson and Shoemaker) who are the object of criticism in his essay (though I do think that Stroud's arguments are sound in this context). Rather, I will consider whether Stroud's argument holds against the neo-Kantian tradition in political philosophy that takes recourse to TAs to defeat the moral nihilist rather than the epistemic sceptic.²⁵

Let me fill in the epistemic-sceptic-as-opponent/moral-nihilist-as-opponent analogy by way of reconstructing a TA that seeks to refute the challenge put to the rational ought-moral ought inference that I argued lay at the core of contemporary neo-Kantian political philosophy. Recall Habermas's response to this challenge. To the nihilist who either (1) refuses the moral commitments implied by communicative action, or (2) refuses to

²⁵In an earlier draft of this paper I paused at this point to defend Kant against Stroud's criticisms. At an analogous point in my consideration of Rorty's criticisms of transcendental arguments I also inserted a defence of Kant against Rorty before my defence of Habermas et al. against Rorty. However, readers of this earlier draft were unanimous in judging these paranthetical sections to be too disruptive to the narrative of my argument. However, I still think that this defence is important, and, as such, have included the passages where I defend Kant against Stroud and Rorty as an appendix to this paper.

engage in communicative action, Habermas responds by claiming (1) that these commitments are the necessary preconditions of communicative action, and (2) that all non-communicative forms of action are derivative of communicative action. To block the nihilist's proposed 'escape route' (e.g., taking on the roles of White's 'first person dictator' or 'systematic free-rider'), Habermas must claim transcendental status for the principles he provides in his response.²⁶ How might such a TA proceed?

Following Genova's analysis, the first step of this TA will be a hypothetical deduction (HD). As with Kant's argument, the HD step of this TA will take the form of a conditional proof. In contrast with Kant's HD, however, this step of the argument will not seek to articulate the "necessary conditions of some contingent domain of experience" (Genova, p. 479), but rather articulate the practical commitments of some contingent domain of action. So rather than taking the form of "given that we perceive objects in the world distinct from the perceiver...", as we might say Kant's HD begins, we will begin with "given that we

²⁶Again, let me emphasize that Habermas would be uneasy about my unqualified use of 'transcendental' to describe the status of the claims of his universal pragmatics. In this respect, my argument has more in common with Apel's position than with Habermas's. Habermas argues that the methodology of the reconstruction of the conditions of speech lies somewhere between the a priori and the empirical. I will take this point up later below. For now, however, let me make it clear that the argument I am providing is something of a 'creative reconstruction.' I am drawing heavily on Habermas's position and insights, and while he would be in agreement with much of what I am saying, I am not claiming textual security as a virtue of my account. That is, my argument ought not to be evaluated in terms of its fidelity with Habermas's work.

engage in communicative action...". Drawing on the material on Habermas that I covered in the last chapter, the argument might proceed as follows²⁷.

- P₁: Speakers of a natural language engage in communicative action, i.e., action oriented to reaching understanding.
- P₂: The act of speaking commits the speaker to raising validity claims concerning the appropriateness and truth-fulness of her or his utterance.
- P₃: 'Being understood' means involves cashing in these claims to validity (or speaking in a context wherein these claims are understood to be satisfied).
- P₄: These validity claims admit of empirical testing.
- P₅: The only criteria available for such testing are those that we devise.
- P₆: These criteria, as well as the particular tests in which they are employed and the investigators who use them, are fallible.
- P₇: The only tribunal in which the adequacy of these tests may be judged is the tribunal formed by other--in fact all other--inquirers.
- P₈: The most positive possible result of such a test is consensus.

Therefore, engaging in communicative action commits us to the desire to reach consensus.

P₉: Given P₃-P₇, moral claims and truth claim have as their final tribunal the potential agreement of all members of the relevant speech community.

²⁷The bulk of this argument is a repetition of elements of the arguments I provided in the past chapter for Habermas's claim that "Our first sentence expresses unequivocally the intention of universal and unconstrained consensus", and his argument for the ideal speech situation. I am repeating this material because, in line with Genova's analysis, it is important to carefully distinguish between the HD and MD levels of argumentation, a distinction overlooked in my initial presentation of the arguments. As well, I suspect readers without photographic memory will appreciate the reminder.

- P₁₀: If agreement reached in critical discussion is to provide a warrant for truth claims, there must be a way to distinguish a rational consensus from a merely de facto consensus.
- P₁₁: We have available to us such a criterion of demarcation: A particular consensus is rational iff it is brought about solely through the arguments employed and not through forces of coercion that are either internal or external to the discussion.

Therefore, engaging in communicative action commits us to attempting to realize a situation of uncoerced dialogue. This commitment involves both behaving a certain way (being willing to cash in the validity claims raised by our utterances), and working towards securing certain social structures (those that will minimize forces that coerce public dialogue). In short, engaging in communicative action commits us to a communicative ethics.

The passage from an HD to an MD essentially involves removing the conditional status of the conclusion of the HD. This, in turn, requires that P₁ no longer take the form of an assumed antecedent. In Kant's terms, we have to demonstrate the 'subjective necessity' of our claim. In this context, that means that we have to show that we cannot refuse to engage in communicative action or--as 'reaching understanding' involves cashing in the validity claims raised in speech--refuse the commitments implied by communicative action.

Translating this part of the argument into syllogistic form is not as instructive as with the HD. The argument seems to precede as follows. Assuming the argument of the HD:

- P₁: All non-communicative forms of action are derivative of communicative action.
- P₂: The refusal to take on the obligations implied by engaging in communicative action *itself* raises the validity claims that (as the HD showed) generate communicative ethics.

Therefore, the structures of our actions and practices commit us to behave in accordance with the obligations of

communicative ethics.

P_1 bars the first escape route out of communicative ethics proposed by the nihilist--choosing to engage in non-communicative action--by subsuming non-communicative action under communicative action. P_2 bars the second escape route proposed by the nihilist--refusing to accept the commitments that flow out of communicative action--by arguing that this refusal itself is a speech act having the characteristics that generate the commitments flowing from communicative action. It follows from P_1 and P_2 that there is no escape route; that the structure of communicative action (and, derivatively, all action) and the communicative ethics that flow from it have the status of subjective necessity.²⁸

P_1 is a piece of conceptual analysis. There is a sense in which its truth follows, though rather trivially, from the truth of P_2 . If we assume P_2 to be true, and define communicative action in terms of the validity claims raised by speech acts--i.e., call communicative action action that raises claim of validity--we could conclude that, insofar as these validity

²⁸ But what status do P_1 and P_2 have? Neither, of course, is an empirical claim. Both, however, claim to be instantiated empirically. This is one of the difficulties that accompanies claiming transcendental (or pseudo-transcendental) status for principles: the claim that certain principles are always instantiated sound like the claim that they are always verified--and at this point the technical problems with Popper's demarcation criterion fade into the background and his intuitions seem sound and precisely relevant here. (Not that TAs propose to be science, of course--but neither do they propose to defend the merely trivially true.) However, I think that P_1 and P_2 are defensible, as I will argue presently.

claims cannot be rejected without being presupposed, then all action must raise these claims, and thus all action must be a form of communicative action. This 'proof' is unsatisfying because it relies on defining communicative action in terms of the validity claims characteristic of communicative action. A more accurate definition of communicative action is 'action oriented towards understanding'--the validity claims raised by speech *follow from* the fact that speech is oriented towards understanding. For P_1 to be successfully defended it has to be shown either that 'reaching understanding' or--as this turns out to be the same thing--'reaching consensus' is somehow the overriding goal of non-communicative action, or that the goals of non-communicative action are derivative of the goals of communicative action. This is a complex and difficult matter.²⁹ We can get a feel for the type of analysis required here if we consider the example of one type of non-communicative action: what Habermas calls strategic action, or action aimed at

²⁹Habermas is not very helpful here. P_1 appeared as a assumption in his earlier writings (Cf., "What is Universal Pragmatics?", p. 1). In his later writings, he worked to clarifying the distinction between communicative and non-communicative action (primarily 'strategic action', i.e., action oriented toward achieving a pre-determined goal) through an analysis of speech acts. (Basically, communicative action is dialogue characterized by an absence of perlocutionary influence, while strategic action is dialogue where the success of speech acts depends on the success of (covertly or overtly) intended perlocutionary forces. Cf. Theory of Communicative Action, vol. 1, pp. 295-205). This analysis, however, remains at the level of the structural differences between communicative and non-communicative action, and does not delve into the deeper type of analysis that I have suggested is necessary to convincingly defend P_1 .

realizing pre-established goals. This is the type of action that much social science (especially economics) and philosophy takes to be the paradigm of rational behaviour. We can get a sense of the claim that such action is derivative of communicative action if we look at some of the shortcomings of the assumption that social behaviour can be explained in terms of agents acting exclusively in accordance with strategic rationality.

Stephen K. White offers the following as a description of this theory of social behaviour.

Action is conceptualized as the intentional, self-interested behaviour of individuals in an objectivated world, that is, one in which objects and other individuals are related to in terms of their possible manipulation. The rationality of action is correspondingly conceptualized as the efficient linking of actions-seen-as-means to the attainment of individual goals.³⁰

This account of rationality runs into problems when it tries to account for cooperative action. The standard line, of course, is that cooperative action is no more or less than the coordination of individual interests. However, this account cannot explain certain forms of behaviour. A classic example is voting. White explains:

[Under this conception of rationality] it appears that a rational individual would decide not to vote, for the simple reason that the cost in time and effort of that act is far too high when measured against the benefit of having his favoured party win, once he considers the likelihood that his one vote will make the difference between his party winning or losing. By not voting, the individual does not measurably change the probability of the collective good (his party in office) being supplied: if his party loses, it would

³⁰White, p. 10.

have done so even if he had not voted; if it wins, he gets the benefit of the collective good without any cost, that is, as a free rider.³¹

White tells us that the response to this problem has been to enlarge the list of motives that the strategic-rational model of social behaviour can include as the objects of rational behaviour. It has been suggested that e.g., "a voter might get some private benefit or satisfaction from *conforming to the norms* of good citizenship or fairness in a democracy."³² I would suggest however, that this marks a departure from, rather than an expansion of, the strategic-rational model. This motivation for voting appears to be a desire to *maintain consensus* about the "norms of good citizenship or fairness in a democracy." As Rousseau pointed out, the institution of majority rule is defensible only on the basis of prior unanimity instantiating the institution. Let us suppose that one of the "norms of good citizenship or fairness in a democracy" that our voter is concerned to protect is the norm of accepting the rule of the majority. I would argue that the agent in this expanded strategic-rational model of social behaviour is doing more than gaining satisfaction. (S)he is participating in maintaining consensus at the prior level of consent that Rousseau argued

³¹White, p. 12.

³²White, p. 13. White gives two references for this response. William Riker and Peter Ordeshook, An Introduction to Positive Political Theory (New Jersey: Prentice Hall, 1973), p. 63; and James Q. Wilson, Political Organization (New York: Basic Books, 1973), chpt. 3.

required unanimity. (S)he is participating, that is, in action aimed at reaching consensus.

Now, this is a very quick argument, and objections may surely be raised to it. But I think that it illustrates the kind of arguments that can be made in favour of the subsumption of strategic under communicative action. In turn, I am suggesting that this illustrates the kind of arguments that can be made for the subsumption of all non-communicative forms of action (not only strategic action, but, e.g, the acts of a 'first person dictator') under communicative action. I do not have the space to develop this argument further. My aim was to illustrate the kind of argument required to defend P_1 , and, by doing so, hopefully make a convincing appeal for its defensibility.

P_2 differs from P_1 in that it does not draw on the type of conceptual analysis that characterized the defence of P_1 . The claim being made by P_2 is more a logical than a conceptual one. The point is that the utterance "I hereby defy the commitment to answer to the validity claims that flows from my utterances" *itself* makes validity claims as to, e.g., the sincerity of the speaker and the truthfulness of the utterance. It follows from this, along with the HD and P_1 , that "the structures of our actions and practices commit us to behave in accordance with the obligations of communicative ethics." This conclusion enjoys the logical status that Genova argued characterized a transcendental principle: that it is implied by its contradictory. As with P_2 , the claim that "the structures of our actions and practices do

not commit use to behave in accordance with the obligations of communicative ethics" *itself* raises the validity claims from which these obligations flow. The conclusion of the MD displays its transcendental status by not being able to be denied without being presupposed.

We have thus moved from an HD--the claim that if we engage in communicative action we are committed to a communicative ethics--to an MD, wherein the hypothetical status of the conclusion of the HD was changed into the categorical by arguing for the inevitability of the assumption that made the first premise of the HD. The next step in the TA is the move from the MD to the TD; the move, that is, from what Kant called 'subjective necessity' to 'objective validity.' It is at this point that Stroud argued a gap in TAs occurred, a gap that must be filled with a VP. The claim I want to defend--and, in the service of which, I have reconstructed this TA--is that no analogous gap and no analogous need appears in a TA of the type I have been constructing on behalf of Habermas, that is, a TA directed, not against the epistemic sceptic, but against a particular kind of moral nihilist.

No analogous gap appears for the same reason that the possibility of an appearance/reality disjunction with respect to our awareness of practical freedom posed no problem for Kant's argument that practical reason can give objective reality to transcendental freedom. Recall from Chapter One that the weakest point in Kant's argument for the objective reality of

transcendental freedom appeared to be the claim that the will can act independently of sensuous impulses. Were this not the case, then pure reason could not be practical; given that it is the case, transcendental freedom exists because practical freedom requires transcendental freedom.³³ Kant argued that practical reason regards with indifference the possibility that our experience of practical freedom is an illusion. Even if in the actions through which it prescribes laws the will is determined at some higher level, this is a question "which in the practical field does not concern us, since we are demanding of reason nothing but the *rule of conduct*" (A 803/B 831). As Beck summarized,

[I]n the practical [realm], the Idea of freedom is constitutive of the experience to which it applies, for the experience is of what ought to be (as defined by the Idea) and not of what happens to exist independently of it.³⁴

In terms that I introduced above, Beck's point--on behalf of Kant--is that the possibility that the realist/sceptic raises for epistemology does not appear here. The realist/sceptic's concern was with the possibility that our epistemic activities do not produce results that match up with the world. But here, the issue of matching up with the world does not arise. So again, we

³³An all too brief summary of my ten page summary of the second Critique, to be sure. But the material was covered in detail in my first chapter, and repetition here would take up unnecessary space. I ask the reader to refer to my account in Chapter One.

³⁴L.W. Beck, A Commentary on Kant's Critique of Practical Reason, p. 48.

take recourse to Kant. Applying Kant's argument for the subordination of theoretical to practical reason--or epistemology to ethics--guarantees that the HD of the above argument assumes that status of a TD. The nihilist with which we are dealing is defeated at the point that the escape routes (s)he proposes are closed with the MD (i.e, at the point at which the MD argued that all forms of non-communicative action are derivative of communicative action, and that the refusal to accept the commitments flowing out of communicative action itself entailed the endorsement of those commitments). As such, while Stroud's arguments may be sound with respect to contemporary neo-Kantian metaphysicians, they do not hold against the manner in which TAs may be employed in the service of (certain forms of) neo-Kantian political philosophy.

Though I think that, within the parameters I have set, the TA above is valid, there are important questions to be raised.

(1) While Stroud's gap does not appear in the course of this TA, it does appear before the argument even begins. Habermas's rational ought-moral ought inference is based on a rational consensus theory of truth. The sceptic is free to point out that such an epistemic theory of truth is always compatible with a metaphysical realist theory of truth. As long as the sceptic's doubts are not conclusively defused (and I doubt they can be), the whole TA above assumes the status of an HD: "Given an epistemological theory of truth..."

(2) I have been careful to argue that this TA defeats a *certain type* of moral nihilist. Is it as effective against other forms of moral nihilism? Consider the position that flat out denies that we can know any rules of conduct, any moral truths. If this is essentially a form of epistemic scepticism, then I think that the position can be met. Habermas and (procedural) ideal limit theorists in general are moral cognitivists, but their meta-ethics does is not committed to any special kind of moral cognition. Their argument is that moral commitments flow out of our practices. At this level, it does not matter if, for reasons familiar to the epistemic sceptic's attacks, we are mistaken about our knowledge of these practices. What we take our practices to be is as good, for these purposes, as what these practices actually are.³⁵

(3) The type of moral nihilism that may pose a more serious threat is the position that argues that no values can follow from facts (and we cannot know values). Again, I have been careful to phrase my position to avoid this attack: I have characterized the inference made in the TA as an ought-ought inference, rather than an is-ought inference. While Habermas grounds his inference in an account of rationality, this account is normative: it is grounded in an account of the distinction between legitimate and binding rather than de facto consensus. His argument (along with Apel and Putnam) is that moral norms flow out of rational norms.

³⁵The point I am making here is analogous to a point that emerged in the recent moral realism debate.

But there is a factual claim underneath Habermas's articulation of the norms of rationality: that speech itself presupposes the desire to reach consensus. I rehearsed the argument for this claim in the last chapter. I think the argument is valid, but hard core defenders of the naturalistic fallacy will disagree.

I do not have the space to meet these objections. What I wanted to establish in this section is that the type of TA required by contemporary neo-Kantians who naturalize Kantian autonomy can meet Stroud's covert verification principle criticism. I will turn now to consider whether this revised TA can meet Rorty's criticisms.

Rorty's criticisms.

In the course of the debate initiated by Stroud, Rorty advances three central claims³⁶: (1) TAs, if legitimate at all, must be restricted in scope; a TA may refute a given alternative conceptual scheme, but cannot legitimize a particular conceptual scheme for all time; (2) TAs as traditionally conceived (i.e., as aspiring to a greater scope than that allowed by (1)) are

³⁶That is, central to my concerns. The thesis Rorty defends in one of these papers is that, while Stroud was correct in claiming that TAs require a VP, the type of verificationism required is not that type for which Stroud argued. Briefly: the brand of verificationism required is not one that depends on a word-world connection, but connections among various parts of our linguistic behaviour. This claim, however, follows from (1); that is, it turns out to be a claim about revised TAs. For my purposes, it is important only to consider Rorty's proposed revision, and assess the soundness of the claims upon which this revision is proposed.

based on an indefensible scheme-content distinction; and (3) ideal limit theory similarly relies on the problematic scheme-content distinction. I will argue, correspondingly, that while (1) Rorty has a strong point about the scope of TAs which draws into questions elements of the TA I presented above; (2) the TA I presented above on behalf of Habermas does not depend on the scheme-content distinction; and (3) ideal limit theory, given the procedural interpretation I have been defending, does not rely on the scheme-content distinction; in fact, it is predicated on the rejection of this distinction. After dealing with these technical criticisms Rorty offers against TAs and ideal limit theory, I will turn, in the next section, to examine Rorty's later, more polemical writings concerning the desire for and feasibility of the transcendental grounding of epistemic and moral claims.

(1) Rorty argues that "the only good 'transcendental' argument is a 'parasitism'" argument."³⁷ A *parasitism* argument shows the sceptic that her conception of the world is parasitic on another conception of the world. If this latter conception of the world turns out to be the one that the sceptic is doubting, then the sceptic's doubts are defused. This is Rorty's interpretation of Strawson's claim in Individuals that I quoted above:

[the sceptic] pretends to accept a conceptual scheme,

³⁷Rorty, "Verificationism and Transcendental Arguments", p. 5. Hereafter, references to this paper will be inserted into my text in the form: (VTA, p.__).

but at the same time quietly rejects one of the conditions of its employment. Thus his doubts are unreal, not simply because they are logically irresolvable doubts, but because they amount to the rejection of the whole conceptual scheme within which alone such doubts make sense.³⁸

Thus, in Rorty's terms, Kant's TA showed Hume that his doubts about causality were *parasitic* upon the employment of the conceptual scheme that Kant defended in the Transcendental Deduction. Rorty generalizes the parasitism argument as follows. The parasitism argument says to the sceptic:

"If you merely say that all the reasons we have for thinking that such-and such's to exist or to be impossible might be insufficient, you cannot be refuted. All that you have done is to say that, in metaphysics as in physics, it is always possible for a better idea to come along which will give a better way of describing the world than in terms of what we thought must necessarily exist, or which will make it possible to recognize the existence of what we previously thought impossible. We can only catch you out if you purport to actually advance such a better idea. Then we may be able to show you that your new way of describing the world would not be intelligible to someone who was not familiar with the old way."
(VTA, p. 5)

Rorty's misgivings come in at the suggestion that a TA--that is, a 'parasitism' argument--could be used to defend a particular conceptual scheme for all time, rather than defeat particular alternatives to a conceptual scheme as they are proposed.

[I]t would be strange if we could know *in advance* of someone's proposing an alternative conceptual framework that it would be parasitic on the conventional one. No one would believe the claim that *any* new theory in physics would necessarily be such that it could never replace, but at most supplement, our present theories. One would have to have an extraordinary faith in the difference between philosophy and science to think that

³⁸P.F. Strawson, Individuals (London, 1959), p. 35.

things could be otherwise in metaphysics. (VTA, p. 11) Rorty makes this point more strongly and more generally in a later paper: "[N]othing in heaven or earth *could* set limits to what we can *in principle* conceive; the best we might do is show that nobody has in fact conceived of an exception [to the received conceptual scheme]."³⁹ So Rorty's claim is that the most a TA can be is a parasitism argument⁴⁰, and no parasitism argument can be offered in advance of proposed alternative conceptual schemes.

This criticism cuts to the heart of the argumentation I presented above in defence of the TA I offered on Habermas's behalf. Especially the first, but also the second premises of the MD depended on the type of predictive power that Rorty denies to parasitism arguments. Recall that P_1 claimed that "all non-communicative forms of action are derivative of communicative action." The argument that I provided as an example of the kinds of arguments that could be generated in defence of P_1 was exactly a parasitism argument. I argued, essentially, that strategic action was parasitic on communicative action, and offered that this type of argument could be generalized into a defence of P_1 .

³⁹Rorty, "Transcendental Arguments, Self-Reference, and Pragmatism", Bieri, et al. (eds.), p. 83. (hereafter TARSP).

⁴⁰In the terms introduced above, for Rorty the only legitimate TA is an HD which takes as its assumed premise some newly proposed alternative conceptual scheme (P) and has as its conclusion a statement of the dependence of P on some already existing conceptual scheme. This turns out to defuse the sceptic if P or an element of P is the refusal of the existing conceptual scheme upon which it is shown that P is parasitic.

Rorty's point would be that even if similar arguments could be constructed for all forms of action we could articulate, there is nothing preventing the introduction of new articulations (or, though this is more difficult to imagine, new types of action). Rorty's point is not as obviously applicable to the defence of P_2 , that is, the claim that "the refusal to take on the obligations implied by engaging in communicative action *itself* raises the validity claims that generate communicative ethics." As I argued above, P_2 , in contrast with P_1 , is more a logical than a conceptual claim. Rorty might argue, however, that this logical claim is grounded in a conceptual claim that sets the boundaries on the possible types of refusal the moral nihilist might advance. Again, Rorty would argue that such an extension is illegitimate. Rorty could claim that the notion of 'performative contradiction' is essentially the notion of 'parasitism.' To argue that refusing the commitments of communicative action involves the speaker in a performative contradiction, Rorty could claim, is the same as arguing that the speaker's refusal is parasitic on the performative structure that (s)he claims to be rejecting. Again, Rorty would say that parasitism could be demonstrated only for particular cases, and that there is no way we could in principle exclude possible future cases.

It should be made clear what Rorty's claim amounts to. According to his analysis, TAs amount to generalized parasitism arguments. Both a Kantian TA and the type of TA required by neo-Kantian political philosophy would thus be essentially inductive

arguments. Now, clearly, TAs (either Kantian or neo-Kantian of the type I proposed above) do not take themselves to be inductive arguments. Rorty's claim, however, is that all that TAs can amount to are (illegitimate) inductively generalized versions of specific parasitism arguments. The problem for the TA that I proposed above on Habermas's behalf is that the arguments for the premises of the MD, especially P_1 , seem to suggest that the premises can indeed only be justified by induction. If that is the case, then Rorty's arguments appear sound. It is difficult to imagine how P_1 could be defended other than by parasitism arguments, and thus it is difficult to offer non-inductive arguments for its universal status. Of course, it would be an illegitimate inductive inference itself to conclude from this argument that therefore no non-inductive argument could be found to defend P_1 and P_2 (and perhaps this is what Rorty himself is doing with respect to the defence of transcendental principles in general). Nonetheless, Rorty succeeds in drawing into question the legitimacy of the universal claims represented by P_1 and P_2 .

(2) Rorty argues that Kantian TAs presuppose a distinction between scheme and content, "e.g., one between concepts and intuitions, or thought and the objects of thought, or words and the world" (TASRP, p. 79). This presupposition is shared with other forms of argumentation; for an argument to be a Kantian TA, it must also satisfy the following conditions:

[a] The scheme-content distinction is construed as a distinction between that which is better known to us (our subjectivity, roughly) and that which is less known to us.

[b] Our 'legitimizing' transcendental knowledge of the necessary truth that content will correspond to scheme, is made possible by the fact that our subjectivity (the scheme) creates the content.⁴¹
(TASRP, p. 79)

Rorty argues that Kant's TA thus becomes problematic when one becomes dubious about the notion of privileged access to one's own subjectivity (TASRP, p. 80).

While Rorty's criticism may hold against the use made of TAs by certain neo-Kantian metaphysicians, I would argue that the type of TA that I have argued is required by neo-Kantian political philosophy neither requires the scheme-content distinction nor depends on the notion of privilege access to one's own subjectivity.⁴² As I argued in response to Stroud's 'verificationism' criticism, the gap that Stroud argued must be filled by a verification principle does not occur in a TA that seeks to defeat a particular kind of moral nihilism rather than epistemic scepticism. The nihilist with which we have been concerned claims that moral oughts do not follow from rational oughts, or that these moral oughts are not binding because the behaviour constrained by the rational oughts may itself be avoided. The goal of a TA here is not to show that a certain way of thinking (a scheme) is legitimate because it fits the world (a content). The nihilist is not defeated on the grounds of an

⁴¹according to Kant's Copernican principle.

⁴²I think, further, that Kant may be defended against Rorty on this point. Please see the appendix for my argument defending Kant against Rorty.

appeal to some legitimized moral cognition. The goal of a TA here is to show that certain commitments follow from engaging in certain types of practices: that engaging in communicative action commits one to a communicative ethics. There is no question here of the correspondence between a scheme and a content.

Similarly, while communicative ethics derives from *our* practices, rather than drawing moral codes from some source outside of our practices, the transcendental legitimation of communicative ethics is not dependent upon privileged access to our own subjectivity--or here, to special insight into our own practices. As I argued above, the sceptical claim that we could be mistaken about our practices along the line of, say, Cartesian global scepticism hypotheses poses no threat to the line of argumentation taken here. For the purposes of this TA, 'apparent practices' are as good as 'real practices.'

In sum, Rorty's claims that TAs depend on a scheme-content distinction and the notion of privileged access to one's own subjectivity are not relevant to the type of TA necessary for neo-Kantian political theory.

(3) Rorty argues that ideal limit theory presupposes the scheme-content distinction. I will argue that, while this may be correct for certain interpretations of ideal limit theory, this claim does not hold for the *procedural* interpretation of ideal limit theory. In fact, the procedural interpretation of ideal limit theory is predicated on a *rejection* of the scheme-content distinction.

The question of whether ideal limit theory depends on the scheme-content distinction depends largely on how you interpret Peirce's claim that "[t]he opinion which is fated to be ultimately agreed to by all who investigate, is what we mean by the truth."⁴³ Traditional interpretations of ideal limit theory (Sellers, Rosenberg, an early version of Putnam) give ideal limit theory what may be called a 'substantive' interpretation. That is, they define the true in terms of the content of the Big Book written at the idealized end of inquiry. As Carl Matheson points out the 'end of enquiry' does not refer to the received view just before the bombs go off, but rather "the state [of inquiry] that we *would* achieve if no external factors were to cut us off prematurely."⁴⁴ What I have called the 'procedural' interpretation of ideal limit theory (hereafter ILT_p) takes Peirce's wording in the oft-quoted phrase above to be rather unfortunate in its use of "the opinion" which is "fated..."⁴⁵

⁴³Peirce, Philosophical Writings of Peirce, p. 38.

⁴⁴Carl Matheson, "Is the Naturalist Really Naturally a Realist?", Mind 48 (1989), p. 253.

⁴⁵As was the case with the position of Rawls's and Habermas's revisions of Kantian autonomy with respect to the Kantian tradition, we are faced with the question of whether the procedural interpretation of ILT marks a continuation of, or departure from the Peircean tradition. I choose to explain the procedural interpretation of ILT as a continuation of the Peircean tradition because (1) Apel, who is the strongest spokesman for the procedural interpretation of ILT , provides a great deal of textual backing grounding his interpretation of Peirce along 'procedural' lines (Cf. Apel, Charles S. Peirce: From Pragmatism to Pragmaticism (Amherst: University of Massachusetts Press, 1981), and (2) because Apel, Putnam and--to a lesser degree--Habermas, who are the major defenders of the procedural interpretation of ILT , all place themselves in the

ILT_p takes the 'limit' to refer to the idealized conditions of inquiry, not the opinion that would be agreed upon under such conditions. As well, ILT_p does not *define* truth in terms of the content of the theories 'fated' to be agreed upon in this idealized state of inquiry; rather, ILT_p argues that if we are compelled to use truth as a predicate, we ought to reserve it to apply to propositions that would be agreed upon under ideal conditions of inquiry. That is, ILT_p resists the implications of the idea of an opinion 'fated to be agreed upon' that suggest that there is a particular opinion--defined counterfactually--that alone can enjoy the status of 'truth.' Rather, ILT_p argues that we reserve our use of the truth predicate for opinions agreed to under the limit of ideal conditions of enquiry, *whatever those opinions may be*. Further, ILT_p makes no commitment to their being a unique opinion-set that may enjoy the compliment 'true.' That is, ILT_p makes no claim to the effect that there is a certain boundary to our knowledge marked by biology or rationality. The procedural limit to inquiry does not imply a unique set of opinions in the way Rawls argues that his

Peircean tradition. That is, I am claiming that it is legitimate for the procedural interpretation of ILT to claim fidelity to Peirce because (1) there is textual backing for this position, and (2) this is the position taken by the prominent spokespersons of this interpretation of ILT. However, if Peirce scholars take exception to Apel's claim to have backing for his interpretation of Peirce, I am prepared to retreat from the textual defence of this position. My position does not stand or fall on this point. My defence of the procedural interpretation of ILT does not need the reference to authority implicit in the claim that this interpretation faithful to Peirce. I think the procedural interpretation of ILT can stand on its own.

principles of justice follow with near-deductive certainty from the original position.

Rorty's misgivings about ideal limit theory draw upon an interpretation of Peirce's statement about truth that gives ideal limit theory what I have called a 'substantive' interpretation. Given this latter interpretation, Rorty correctly questions the attribution of privileged status to the opinion which would be agreed upon at the end of inquiry.

Peirce could have left out the eschatology and just said that truth consisted in that theory which best predicted and accounted for everything which one could ever wish for--where 'best' here means "optimizing all the usual desiderata of theories--elegance, simplicity, familiarity, and the like." (TASRP, p. 84)

ILT_p agrees with Rorty that 'true' ought to be used to describe that which "accounts for everything which one could wish for", but insists--against Rorty--that even *this* use of 'true' posits an ideal limit. This latter ideal limit is the idealized state of inquiry; in Habermas's terms, the 'ideal speech situation.' ILT_p begins with the claim that the only criteria we have of adjudicating over the judgements and claims of fellow-inquirers are the criteria that we devise. It follows from this (through an argument which has been rehearsed twice above) that, given the fallibility of the application of these criteria, the final tribunal by which the application of 'true' may be adjudicated is the tribunal formed by all possible participants of inquiry. The most positive result of such an inquiry is consensus--not de facto consensus, but consensus formed under conditions free from

coercive influences on debate. These latter conditions form the ideal limit of inquiry. So even Rorty's reserved use of 'true' posits an ideal limit, but a different kind of ideal limit from the interpretation of ILT which he attacks.

The traditional interpretation of ideal limit theory (namely, that given by Sellers, Rosenberg, and an earlier Putnam) supposes that there is something metaphysically special about the content of the opinion 'fated' to be agreed upon at the end of inquiry. For example, Rosenberg argues that

because a language is the evolutionary product of the continuing causal impact of anomolous experiential inputs to the community of representers that the structure of its *extensional* realization tends increasingly towards proto-correlational isomorphism with the world represented.⁴⁶

Rorty correctly argues that such a claim is indefensible. There is no reason to believe that the more scientific theories satisfy our criteria for good theories that they are approaching the (metaphysically distinct) truth. There is no reason to believe, that is, that the more our scheme satisfy the criteria by which we judge this scheme, the closer it is to accurately depicting the content which it seeks to describe.

But ILT_p does not require such a scheme-content distinction. In fact, it is predicated on a rejection of this distinction. ILT_p asks that we abandon our talk of 'truth-as-correspondence' and replace it with talk of truth as 'ideal assertability' or

⁴⁶Jay Rosenberg, Linguistic Representation (Dordrecht, 1974), p.119. Rorty's citation.

'that which is the object of rational consensus.' No mention is made in this conception of the metaphysical status of the propositions to which we are justified in attaching the proposition 'is true.' ILT_p asks us to abandon our talk of truth as correspondence because this interpretation of ideal limit theory maintains that the final tribunal of testing truth claims is the consensus of inquirers who, *under uncoerced conditions*, assent both to the criteria proposed for such evaluation and the acceptability of a particular instance of the employment of these criteria. That is, ILT_p begins by rejecting the demand that 'truth' be used to describe a matter of the correspondence of an explanatory scheme with the content of that which is to be explained.

So while Rorty's claim that ideal limit theory requires the scheme-content distinction does apply to traditional interpretations of ideal limit theory, it does not apply to the procedural interpretation of ideal limit theory. In fact, ILT_p is predicated on the rejection of the scheme-content distinction.

Rorty and neo-Kantian political philosophy

In this section, as I promised in the introduction to this chapter, I will step back from the more technical issues concerning transcendental justification which dominated the last section, and move on to the more general problems of neo-Kantian political philosophy raised by Rorty. I accept much of Rorty's claims, but argue that his central critical claims do not have

the implications that Rorty argues for (or, alternatively, that the general picture Rorty asks us to accept is not supported by the critical claims he advances in the service of this defence). Rorty's critique of neo-Kantian political philosophy will help me make explicit what has remained an implicit preference for the Habermas-Apel-ILT_p current of neo-Kantianism over the Rawlsian revisions of Kant⁴⁷. My defence of this partisanship within contemporary neo-Kantian political philosophy will take us back to differences between Rawls and Habermas that I explicated in the past chapter in the context of Rawls's and Habermas's relative fidelity to Kant.

My most fundamental difference with Rorty's picture of liberal politics without transcendental justification is a generalization of a point I raised at the end of the last section: that, *pace* Rorty, using 'knowledge' or 'truth' "simply as a compliment paid to the beliefs which we think so well justified that, for the moment, further justification is not needed"⁴⁸ commits us to an ideal limit; an ideal limit conceived, however, not in terms of content, but in terms of procedure.

⁴⁷An implicit preference insofar as it has been evidenced only by the exclusion of Rawls from the proceeding discussion.

⁴⁸Rorty, "Solidarity or Objectivity?", John Rajchman and Cornel West (eds.), Post-Analytic Philosophy (New York: Columbia University Press, 1985), p. 7. Rorty takes this 'compliment' theory of truth to be the heart of limit-free pragmatism. I will argue, however, that it is this precisely this account of truth that requires us to accept the limit posited by ILT_p.

Rorty has called his positive political position "postmodern bourgeois liberalism"⁴⁹. Postmodern bourgeois liberalism, Rorty tells us, is an attempt to defend the institutions of the rich North American democracies without "traditional Kantian buttresses, buttresses which include an account of 'rationality' and 'morality' as transcultural and ahistorical."⁵⁰ 'Postmodern bourgeois liberalism' is 'bourgeois' because it has no quarrel "with the Marxist claim that a lot of those institutions and practices [which postmodern bourgeois liberalism defends] are possible and justifiable only in certain historical, and especially economic, conditions." (PBL, p. 216). 'Postmodern bourgeois liberalism' is 'postmodern' in the sense given to this term by Jean-Francois Lyotard. In The Postmodern Condition, Lyotard defined 'postmodern' as "incredulity toward metanarratives."⁵¹ Lyotard offers as examples of metanarratives

⁴⁹He later called it "liberal ironism", but these terms are, by Rorty's account, essentially equivalent. An 'ironist' is "the sort of person who faces up to the contingency of his or her own most central beliefs and desires--someone sufficiently historicist and nominalist to have abandoned the idea that those central beliefs and desires refer back to something beyond the reach of time and chance." (Rorty, Contingency, Irony and Solidarity (Cambridge: Cambridge University Press, 1989), p. xv.). As we will see, a postmodern bourgeois liberal is basically a North American liberal ironist, in Rorty's terminology.

⁵⁰Rorty, "Postmodern Bourgeois Liberalism", Robert Hollinger (ed.), Hermeneutics and Praxis (Notre Dame: University of Notre Dame Press, 1985), p. 216. Hereafter, references to this paper will be inserted into my text in the form (PBL, p. __).

⁵¹Jean-Francois Lyotard, The Postmodern Condition: A Report on Knowledge (Minneapolis: University of Minnesota Press, 1984), p. xxiv.

"the dialectics of Spirit, the hermeneutics of meaning, the emancipation of the rational or working subject, or the creation of wealth."⁵² Postmodern bourgeois liberals direct their incredulity specifically toward Kantian principles which offer justifications for liberal institutions. Among these Kantian principles, postmodern bourgeois liberalism is most concerned to abandon the formal justifications that make reference to a meta-community--humanity--to which our loyalty ought to be directed, rather than simply toward the community in which we are members. So, against Kantian attempts to justify liberal ideals like equality, Rorty offers a 'naturalized Hegelian' defence of loyalty only to specific, contingent communities, a defence predicated on the rejection of the Kantian predisposition to the ahistorical. Rorty's most recent book, Contingency, Irony, and Solidarity, is largely given to filling in this naturalized Hegelian defence of liberalism.

"The crucial move in this reinterpretation [of the defence of liberal values]", Rorty tells us,

is to think of the moral self, the embodiment of rationality, not as one of Rawls's original choosers, somebody who can distinguish her *self* from her talents and interests and views about the good, but as a network of beliefs, desires, and emotions with nothing behind it--no substrate behind the attributes." (PBL, p. 217).

Rorty's argument here is not ontological, but methodological:

For the purposes of moral and political deliberation and conversation, a person just *is* that network, as for purposes of ballistics she is a point-mass, or for

⁵²Lyotard, p. xxiii.

purposes of chemistry a linkage of molecules.
(PBL, p. 217)

I think Rorty's anti-essentialism (or, at least, methodological anti-essentialism) with respect to person-hood is on the mark.⁵³ I do not want to commit myself to an ontological position with respect to personal identity here. Along with Rorty, I do not think I have to. To put the point in terms familiar from similar points made above, for the purposes of political theorizing, 'apparent persons' are as good as 'real persons.' That is, I am arguing that even if we could come up with the definitive argument establishing once and for all the ontological status of personhood (and I cannot imagine how such a claim is possible), it would have no implications for political theory. The relevant considerations for political theory are how we understand ourselves in relation to the persons and institutions around us-- and I think that Rorty⁵⁴ is correct in suggesting that the proper language of self-description (at least in this context) is the language of beliefs, desires, and emotions, rather than the

⁵³However, the reasons why I agree with Rorty's position do not derive as much from Rorty's arguments for the 'contingency of selfhood', as from Foucault's arguments about the normative role of the human sciences in the constitution of what we take personhood to consist in. Foucault's arguments will be taken up in detail in the final chapter.

⁵⁴Along with Charles Taylor, Cf. Taylor, "Self Interpreting Animals", Human Agency and Language: Philosophical Papers One (Cambridge: Cambridge University Press, 1985), pp. 45-76.

language of the ontological status of personhood.⁵⁵

My preference for the Habermas-Apel current within contemporary neo-Kantian political theory over the Rawlsian alternative derives most fundamentally from considerations like those concerning personhood advanced above. Though Rawls's theory of justice is not grounded in a strongly essentialist theory of personhood, Rawls certainly supposes (and must suppose--given the structure of his naturalized Kantianism) a great deal about universal human interests. Recall the features of the veil of ignorance. While agents behind the veil are deprived of knowledge of their place in society, class position or social status, fortunes in the distribution of natural assets and abilities, and conceptions of the good or special psychological propensities (TOJ, p. 12), they do know the general facts about human society, including a grasp of politics, economics and psychology (TOJ, p. 137), and they are allowed to assume a preference for more basic social goods rather than fewer (TOJ, p. 142). Thus Rawls excludes what he--along with a major current of the liberal tradition--takes to be morally arbitrary within

⁵⁵The alternative route for grounding moral or political theories in accounts to personhood is natural law. Though I do not hold that the is-ought gap cannot be crossed, I think the natural law route is illegitimate.

Let me make clear that I do not intend by my claim to rule out theories like Hobbes's. I think that Hobbes's justification for his political state can be run without reference to human nature. Rather than claiming that humans are, by nature, ethical egoists (as some readings of Hobbes impute to him), all we need to say is that there exists the possibility that at least one person may act in accordance with ethical egoism. This is enough to start us down the very slippery Hobbesian slope.

personhood, but he includes much of what he takes to be essential to personhood. (In fact, we could read the veil of ignorance as a mechanism to access the essential--perhaps *morally* essential--elements of personhood). The 'general facts of society' clearly include general facts of personhood that Rawls takes to be non-controversial; psychology clearly makes claims about human nature, most economic theories are underscored by assumptions about human behaviour, as are models of political change and stability.⁵⁶ The assumption that persons want more rather than fewer basic social goods--and the assumption that there is a non-controversial list of basic social goods--has embedded in it something of an essentialism with respect to personhood, as well.

Habermas (and ILT_p in general) makes no such assumptions. The difference between Rawls and Habermas on this point is essentially the difference between Rawls and Habermas that I pointed to in the last chapter when I argued that Habermas's revision of Kant is more purely procedural than Rawls's. For Rawls, consensus is deviant when it is informed by morally arbitrary facts. For Habermas, and what I have been calling ILT_p in general, consensus is deviant when the process of consensus-formation *itself* is distorted. Habermas and ILT_p , by characterizing rational consensus in purely procedural terms, are not committed to any kind of essentialism with respect to

⁵⁶Embedded in Rawls's characterization of the basic facts of society is a characterization of social science as modelled on natural science. This is another major point of difference I have with Rawls.

personhood, as Rawls is (as evidenced by the dependence of Rawls's account of rational consensus on a particular conception of the morally arbitrary characteristics of persons).

Let me review the argument so far. I have argued that Rorty's position serves to question the Rawlsian current of contemporary neo-Kantian political theory because Rorty correctly casts doubt on the appropriateness of grounding political theory in assumptions concerning human nature. As I noted above, Rorty's anti-essentialist leanings in this area are methodologically rather than ontologically motivated. I argued, on Rorty's behalf, that all we have to (and can) deal with in political theory is what persons take themselves to be, and, in this context, the language of self-description is limited to matters like desires and beliefs. The stronger point--which I cannot properly defend here--is that desires and beliefs themselves are importantly informed by the political process. Here I begin to depart from Rorty, as my intuitions at this point are generally informed by Marxism; the concept of 'false consciousness' is lurking around the corner. But such a concept need not be invoked, at least not in its vulgar form. We can turn to Habermas. The analogue in Habermas for the orthodox notion of 'false consciousness' is a set of beliefs, desires, etc. that are formed in the context of a public domain wherein the structures of communication are distorted in a manner that prevents the proper testing of the universal status of normative claims (i.e., a public domain where the structures of

communication allow the interests of some to be represented as the interests of all). I think it is theoretically respectable to speak in these terms, as I will defend further below. Speaking this way suggests limitations to a political theory that grounds itself in any pre-political notion of beliefs and desires. Rawls's theory does just that, as evidence by the constraints Rawls places on agents behind the veil of ignorance.

Rorty's anti-essentialist leanings are not limited to the issue of personhood, however. Rorty argues that pragmatism (James's and Dewey's but not Peirce's) demands, among other things, that we abandon not only theories of persons, but theories of all other things about which philosophers have traditionally theorized. Insofar as Rorty characterizes his project as carrying out the implications of James's and especially Dewey's writings, it will be instructive to investigate Rorty's characterization of pragmatism. Rorty suggests that there are three basic claims central to this form of pragmatism. He argues, further, that these claim rule out not only realist theories of truth and justification, but ideal limit accounts of truth and justification. I will argue that, given the procedural interpretation, ideal limit theory is not only compatible with Rorty's position as he explains it via the central claims of Deweyian pragmatism, but that the idea of an

ideal procedural limit is required by such a pragmatism.⁵⁷ The position I am arguing here is little more than a generalization of the point that I made at the close of the last section of this chapter: that, *pace* Rorty, the 'compliment' theory of truth (i.e., the claim that 'true' is a compliment that we pay to propositions that we feel justified in asserting) commits us to an ideal limit--but a procedural rather than substantive limit.

I will consider the three characteristics Rorty advances on behalf of pragmatism separately, and then turn to a more general comparison of Rorty's pragmatism/postmodern bourgeois liberalism and ILT_p.

§1 "[Pragmatism] is simply anti-essentialism applied to notions like 'truth,' 'knowledge,' 'language,' 'morality' and similar notions of philosophical theorizing."⁵⁸ As usual, Rorty's enemy here is the notion of truth as correspondence to reality. He illustrates what he means by anti-essentialism by reference to James's definition of 'the true' as 'what is good in

⁵⁷My position thus amounts to the claim that if you want to be a pragmatist, you have to follow the route that Peirce originally charted. In this context, I am defending ideal limit theory against Rorty by defending Peirce against Dewey, given that Rorty draws heavily on Dewey, and Habermas and especially Apel draw heavily on Peirce. This would be a tidy way to make my point, but my acquaintance with Peirce is very basic, and I have read little of Dewey's writing. I think my intuitions on this point are correct. But I have to leave this claim for a later project.

⁵⁸Rorty, "Pragmatism, Relativism, and Irrationalism", Consequences of Pragmatism (Minneapolis: University of Minnesota Press, 1982), p. 162. Hereafter, all references to this paper will be inserted into my text in the form: (PRI, p.____).

the way of belief.' (PRI, p. 162). While the practice of pairing off bits of language with bits of the world may be a useful exercise when we are dealing with sentences like "This is water," it seems inappropriate, Rorty argues, to use the 'mapping' metaphor when we get to the level of theory, e.g., when we ask "what exactly it is that Marx pictured more accurately than Machiavelli." (PRI, p. 163). James's point, Rorty tells us, is that carrying out the mapping exercise "will not enlighten us about why truths are good to believe, or offer any clues as to why or whether our present view of the world is, roughly, the one we should hold...[I]t is the vocabulary of practice rather than of theory, of action rather than contemplation, in which one can say something useful about truth." (PRI, p. 162). Rorty argues that this view amounts to (but, of course, is not equivalent to) the claim that there are "no essences in the area" of truth, rationality, inquiry, or the relation of thought and object.

All of this is compatible with ILT_p . All that ILT_p would want to add to this picture is the claim that our reflections about which truths are good to believe or "why or whether our present view of the world is, roughly, the one we should hold" take place within the arena of the public exchange of ideas, and that this process of debate admits of, and requires, criteria to distinguish between rational and merely de facto consensus. ILT_p does not claim that the essence of truth is 'rational consensus' any more than James's claims that the essence of truth is "what is good in the way of belief." We might say that ILT_p agrees

that we award 'true' to that which a community finds to be "good in the way of belief." But ILT_p argues that such a conception of truth must be saved from becoming arbitrary or, worse, a cloak for the exercise of power, and that this can be done by ensuring that the standards and procedures adopted in forming agreement are non-coercive.

It may be countered that even if ILT_p does not harbour a covert essentialism with respect to truth, it certainly requires a linguistic essentialism. Habermas's argument for communicative ethics required making strong claims about the necessary conditions for communicative action. Isn't this a claim about the essential characteristics of language?

Perhaps there is something to this argument. While ILT_p 's analysis of speech does not embody the same type of essentialism with respect to language as, say, the writings on language characteristic of the hermeneutic tradition, ILT_p is clearly grounded in an analysis of the *nature* of speech and (successful) communication. While the validity claims that Habermas argues accompany the act of speaking can be empirically tested, the claim that speech raised validity claims *itself* admits of no empirical vindication. In this sense, Habermas's analysis of speech and communication does seem to be saying something about the nature of speech, and, thus, in this sense, ILT_p is incompatible with what Rorty argues to be one of the basic characteristics of (non-Peircean) pragmatism. However, as I will argue below, I think that Rorty's pragmatism is (or, perhaps,

ought to be) committed to the type of limit talk that characterizes ILT_p . If such limit-talk must involve the type of linguistic essentialism involved in Habermas's analysis of speech, then pragmatism must concede that such a limited essentialism is necessary.⁵⁹

§2 "[T]here is no epistemological difference between truth about what ought to be and truth about what is, nor any metaphysical difference between facts and values, not any methodological difference between morality and science." (PRI, p. 163). Point for point, ILT_p is in agreement with this claim. ILT^p would agree that "there is no epistemological difference between truth about what ought to be and truth about what is" because, according to ILT_p , claims about what is and claims about what ought to be share the same tribunal: the scrutiny of inquiry and rational consensus-formation by the relevant speech community. Descriptive and normative claims both stand on the same epistemic footing. ILT_p would agree that "[there is] no metaphysical difference between facts and value" in the trivial sense that ILT_p is a purely epistemic account of truth and value. Perhaps the more accurate claim, then, would be that ILT_p is *agnostic* with respect to the metaphysical status of facts and values, or perhaps indifferent to the issue. This would be a

⁵⁹Again, this point could be made succinctly by arguing that Dewey's critique of Peirce was wrong in claiming that the ideal limit was an idea unnecessarily tacked onto pragmatism. But I have to leave this claim as a hypothesis to be investigated in a later project.

more accurate statement of Rorty's position, as well. Rorty is at pains to emphasize that pragmatism does not offer theories of philosophical problems, in the tradition sense of philosophical theories. As far as this goes, then, Rorty cannot claim that there is no metaphysical difference between facts and value; this itself is a metaphysical claim.⁶⁰ Finally, ILT_p would agree that "[there is no] methodological difference between morality and science" inasmuch as ILT_p places descriptive and normative claims on the same epistemic footing. I am not claiming (on behalf of ILT_p) that scientific and ethical methodology is the same at the level of day-to-day practice: arguing about, say, the abortion issue and testing a physical hypothesis are methodologically distinct practices. The level of methodology that I suspect Rorty is aiming at here is the same meta-level at which ILT_p claims there is a continuity in method between morality and science, i.e., the level at which the final test of the acceptability of a moral or scientific claim is its being the object of rational consensus. This leads us into Rorty's final characterization of pragmatism.⁶¹

⁶⁰More importantly, Rorty's position is completely compatible with any metaphysical position. Even when Rorty asks us to abandon metaphysical theories of truth, it is because he feels that such accounts of truth are useless, rather than incorrect.

⁶¹Putnam offers a succinct statement about the blurring effect ILT has on the fact/value distinction. Echoing Habermas's claim that "the truth of statements is based on the anticipation of the realization of the good life," Putnam argues that "[t]he notion of truth itself depends for its content on our standards of rational acceptability, and these in turn rest on and presuppose our values. Put schematically and too briefly, I

§3 "[P]ragmatism is the doctrine that there are no constraints on inquiry save conversational ones--no wholesale constraints derived from the nature of objects, or of the mind, or of language, but only those retain constraints provided by our fellow inquirers." (PRI, p. 165). ILT_p would add to the above claim the claim that we can and must distinguish among legitimate and illegitimate (or admissible and inadmissible) constraints that our fellow inquirers may provide us. Here is where my argument shifts from claiming that Rorty's characterizations of non-Peircean pragmatism are *compatible with* ILT_p to claiming that Rorty's pragmatism must commit itself to the type of ideal limit for which ILT_p argues.

Rorty's argument against limit-constraints on inquiry echoes his argument about the restriction of the scope of TAs to parasitism arguments advanced against particular alternative conceptual scheme. Rorty argues (and echoes Kant's Copernican principle in doing so) that the pragmatist claims that "it is useless to hope that objects will constrain us to believe the truth about them." (PRI, p. 165). "The only sense in which we are constrained to truth", Rorty continues, "is that, as Peirce suggested, we can make no sense of the notion that the view which can survive all objections might be false." (PRI, p. 165). But Rorty leaves Peirce and limit-talk behind at this point.

am saying that theory of truth presupposes theory of rationality which in turn presupposes our theory of the good."

Putnam, Reason, Truth, and History (Cambridge: Cambridge University Press, 1981), p. 215.

But objections--conversational constraints--cannot be anticipated.⁶² There is no method for knowing *when* one has reached the truth, or when one is closer than before. (PRI, pp. 165-166)

As above, when I defended ILT_p against Rorty's claim that ideal limit theory was committed to a scheme-content distinction, I think Rorty's argument here is restricted only to the interpretation of the ideal limit that awards a metaphysically privileged position to the opinions reached "at the end of inquiry." As I argued above, ILT_p makes no such claims, and is in fact predicated in a rejection of the scheme-content distinction that informs such a metaphysical reward. ILT_p shares with Rorty his 'compliment' theory of truth, i.e., the claim that the truth predicate is awarded as a compliment to propositions that our evaluational criteria deem warranted to assert⁶³. ILT_p departs from Rorty's pragmatism with ILT_p 's claim that the 'truth compliment' is paid to propositions that pass the *legitimate* or *admissible* constraints that our fellow inquirers may put on inquiry. Rorty shuns such a de facto/de jure distinction.

But Rorty also speaks as though such a distinction is relevant. In Contingency, Irony, and Solidarity, Rorty offers as one characterization of his post-Kantian liberal society: "A liberal society is one which is content to call "true" (or

⁶²Here is where the parallel with Rorty's critique of TAs occurs.

⁶³rather than awarding 'true' as a compliment paid to propositions that enjoy some metaphysically privileged status, e.g., propositions whose content corresponds to the world.

"right" or "just") whatever the outcome of *undistorted* communication happens to be, whatever view wins in a *free and open encounter*."⁶⁴ Earlier in the book, within the context of distinguishing between reasons for belief and causes for belief, Rorty argues that we can draw a distinction between persuasion and force by generalizing from obvious examples, like "that between Socratic dialogue and hypnotic suggestion." He continues:

We then try to firm up the distinction by dealing with messier cases: brainwashing, media hype, and what Marxists call "false consciousness." There is, to be sure, no neat way to draw the line between persuasion and force, and therefore no neat way to draw a line between a cause of changed belief which was also a reason and one which was a "mere" cause. But the distinction is no fuzzier than most. (CIS, p. 48)

Rorty's solution to the problem that certain agreements seem worthy of greater respect than others is to take recourse to generalizing from paradigm examples. In "Pragmatism, Relativism, and Irrationalism", he offers a solution similar to that from CIS above. Recognizing that the idea of undistorted communication is relevant to the pragmatist account of truth and goodness, Rorty nonetheless leaves Habermas when Habermas "goes transcendental and offers principles." Against this approach, Rorty claims that "[t]he pragmatist...must remain ethnocentric and offer examples."

He [the pragmatist] can only say "undistorted" means employing *our* criteria of relevance where we are the people who have read and pondered Plato, Newton, Kant,

⁶⁴Rorty, Contingency, Irony, and Solidarity (Cambridge: Cambridge University Press, 1989), p. 67). (my emphasis) Hereafter, all my references to this book will be inserted into my text in the form: (CIS, p. __).

Marx, Darwin, Freud, Dewey, etc. Milton's "free and open encounter," in which truth is bound to prevail, must itself be described in terms of examples rather than principle--it is to be more like the Athenian market-place than the council-chamber of the Great King, more like the twentieth century than the twelfth, more like the Prussian Academy in 1925 than in 1935. (PRI, p. 173)

So, in this context, Rorty's pragmatism differs from ILT_p in terms of the means by which undistorted communication can be specified--for Rorty, the criterion is by generalized example, for ILT_p , the criterion is procedural. Rorty regards recourse to a procedural account of undistorted communication as an unjustifiable recourse to the transcendental, an illegitimate abstraction from contingent practices. ILT_p regards Rorty's reliance on examples from history as uncritical and potentially ideological, an illegitimate generalization of contingent practices.

We appear to have a version of the problem of the criterion here. Rorty would say to ILT_p that any principle ILT_p would offer to distinguish undistorted from distorted communication would be no more than a generalization of examples of paradigmatic undistorted communication. ILT_p would ask Rorty how he identified these examples as paradigmatic of undistorted communication. ILT_p has somewhere to go to justify itself: the transcendental conditions of communication. Rorty has nowhere to go, but is content to remain 'on the surface' because of his commitment to the pragmatic injunction against the transcendental and the a priori. Can Rorty have his cake--and take no recourse

to the transcendental--and eat it too--but still be able to meaningfully distinguish between distorted and undistorted communication?

If this is really an occurrence of the problem of the criterion, then there is no a priori way to adjudicate over the debate between the procedural vs. 'generalized example' criterion for demarcation. However, while such a debate is afforded no a priori resolution, it seems that convincing arguments can be offered in favour of one side. Rorty's arguments for his position come from his appeal for the pragmatist no-theory theories of 'truth', 'goodness' and the like. ILT_p defends itself by arguing that there is nothing illegitimate or metaphysically troublesome about admitting that certain of our practices seem to posit idealized versions of themselves. ILT_p asks us to recognize an implicit criterion in our predisposition to accord more authority or authenticity to an agreement made in the course of a friendly conversation than an agreement made between a person holding a loaded gun and the person at whom the gun is pointed. The latter is clearly a situation of coercion; ILT_p asks us to recognize that our suspicion with respect to the outcome of deliberation in such a situation is informed by an implicit condition in speech that demands that consensus is 'rational' iff the process by which it is secured is free from coercion.

Putnam has put his finger on the balance the ILT_p attempts to strike.

There are two points that must be balanced, both points that have been made by philosophers of many different kinds: (1) talk of what is "right" and "wrong" in any area only makes sense against the background of an inherited tradition; but (2) traditions themselves can be criticized.⁶⁵

The advantage accrued by adding the qualification of (2) above is exactly the advantage gained by following the procedural path in distinguishing undistorted from distorted communication. A procedural distinction between distorted and undistorted communication affords us critical space in which to ground our judgments prompted by the suspicion that even the 'paradigm' cases of undistorted communication may harbour a coercive undercurrent. In short, and in traditional terminology, the procedural approach to demarcation allows us to ground the critique of ideology.

Rorty is correct in suggesting that such a critical project is predicated on an appeal to something of an extra-historical community. At this level, ILT_p is directly opposed to Rorty's claim that all we can talk about and all that we need to feel is loyalty toward the contingent historical community of which we are members. But Rorty has constructed something of a strawman for himself. The alternative to his view need not be the "traditional Kantian buttress" of reference to the "supercommunity" of "humanity as such" (PBL, p. 214). ILT_p is indeed committed to the notion of membership in a community

⁶⁵Putnam, "Why Reason Can't Be Naturalized", Kenneth Baynes, James Bowman, and Thomas McCarthy (eds.), After Philosophy (Cambridge: MIT Press, 1987), p. 227.

parallel to the contingent historical community, but the parallel community of which ILT, speaks is not grounded in any essentialist notion of personhood. In Apel's words,

anyone who engages in argument automatically presupposes two things: first, a *real communication community* whose member he has become through a process of socialization, and second, an *ideal communication community* that would basically be capable of adequately understanding the meaning of his arguments and judging their truth in a definitive manner.⁶⁶

Apel argues that an *ideal communication community* is posited by argumentation; his reasons draw on Peirce and will be familiar from my discussion of Habermas's argument for the ideal speech situation. Given that any particular application of any evaluative criterion is fallible, and that evaluative criteria are always open to revision, no answer to a given test is conclusive. A proposition may be rewarded the status of 'true', ideally, if all possible tests were run on it, and if all possible inquirers were satisfied by these tests. Here we turn Rorty's arguments against ILT against him. Rorty argued that a 'limit' is meaningless because we could never tell when we got to the truth. It is for this very reason that Apel argues for an ideal communication community. It is precisely because the evaluative procedure cannot be realized in the real communication community that we recognize the idealized status of our practice of complimenting proposition with 'truth.' The fallibility of our evaluative procedures points toward the standard by which our

⁶⁶Karl-Otto Apel, Towards a Transformation of Philosophy (London: Routledge & Kegan Paul, 1980), p. 280.

tests are to be judged, and this standard is an idealized version of the imperfect community of inquirers of which we are members. This disjunction, the real communication community/ideal communication community--like the kingdoms of the earth/kingdom of ends disjunction in Kant--affords ILT_p , a grounding for the critique of the institutions and practices that, to a lesser or greater extent, allow our societies to approximate the ideal communication community.⁶⁷

The upshot of my argument is this: insofar as the debate between Rorty and ILT_p over the means by which to distinguish between distorted and undistorted communication is a version of the problem of the criterion, then there are no a priori grounds for preferring one over the other. But ILT_p offers more than Rorty's alternative. ILT_p offers a critical space from which to build a critique of ideology. In Putnam's terms, ILT_p agrees with Rorty that "talk of what is 'right' and 'wrong' in any area only makes sense against the background of an inherited tradition", but argues, against Rorty, that "traditions themselves can be criticized." Further, ILT_p offers this normative edge at a lesser cost than Rorty's estimation. ILT_p 's neo-Kantian 'buttresses', I would argue, are not nearly as philosophically contentious as the traditional Kantian claims. ILT_p arrives at the idea of a limit from the same point from

⁶⁷"The contradiction that must be overcome between the real and ideal communication community...supplies the starting-point for an evaluative critique of ideology." Apel, p. 145.

which Rorty argues that we abandon the idea of a limit: the recognition of the fallibility and inconclusiveness of specific, historical evaluations of hypotheses. Rather than abandoning the idea of an ideal, ILT_p argues, along with Peirce, that arriving at the notion of, say, the 'real', involves us in committing ourselves to hypothesis testing that is "independent of the vagaries of me and you,"⁶⁸ and requires "a conceived identification of one's interests with those of an unlimited community."⁶⁹ An unlimited community, while a necessary concept given the fallibility and inconclusiveness of specific, historical evaluations of hypotheses, is nonetheless an unrealizable ideal. But it functions precisely as an ideal; a normative ideal against which our contingent, specific descriptive and normative claims may be judged, and an ideal towards which our political critiques and concrete proposals for reform may be directed.

So, in sum, I have argued that

(1) The current within contemporary neo-Kantian political theory represented by Habermas and Apel (and, to a lesser degree, Putnam)--which I have collected under the idea of an procedural interpretation of ideal limit theory--successfully naturalizes

⁶⁸Peirce, "Some Consequences of Four Incapacities," Justis Buchler (ed.), Philosophical Writings of Peirce (New York: Dover, 1955), p. 247.

⁶⁹Peirce, "On the Doctrine of Chances, With Later Reflections," Buchler (ed.), p. 163.

Kant's account of autonomy while retaining Kant's account of the normative.

(2) While this naturalization necessitated taking recourse to the transcendental, ILT_p revises TAs in a way that escapes the most important critiques of TAs in contemporary literature.

(3) The current of neo-Kantian political theory represented by ILT_p escapes Rorty's most important critiques of the Kantian tradition in political theory, and makes a convincing case for the adoption of a procedural limit as a constraint on inquiry. ILT_p argues that, *pace* Rorty, construing 'truth' as a compliment paid to propositions we feel justified in asserting requires us to posit a idealized version of the process of inquiry.

The most significant challenge to contemporary neo-Kantian political theory comes from those who attack the neo-Kantian justification of political critique via challenging the neo-Kantian notion of autonomy at the practical, political level. The strongest spokespersons of this critique are Foucault and a current of feminist critique of liberalism represented here by Iris Marion Young. It is to these criticisms that I turn in my final chapter.

CHAPTER FOUR**AUTONOMY AND PERSONHOOD**

My general argument thus far has focused on the articulation and defence of the neo-Kantian current in contemporary political theory that I have collected under a position which I have called a procedural interpretation of ideal limit theory. In the first chapter I argued that we can read Kant as offering a solution to the general problem of the relationship between the methodological and substantive concerns of normative political theory. Specifically, Kant's answer was developed in the form of a link between the question of the foundation of political critique and the question of the conditions of autonomy. For Kant, and the neo-Kantian current of contemporary political theory, autonomy is understood as an issue which appears at the cross-roads of both philosophical and political questions. The most distinctive elements in Kantian tradition in political theory are the assignment of centrality to the role of autonomy in normative political theory, and the recognition that 'autonomy' has both transcendental and political components, neither of which is reducible to the other. To review: the issue surrounding the transcendental precondition of autonomy is basically the metaphysical issue of free will, i.e., the question of whether our experience of acting freely is an illusion, or if it matters whether it is an illusion. The issue of the political preconditions of autonomy is the question of what practical

conditions must obtain to allow the free will to be exercised in the public realm.

Kant's account of autonomy, however, contained as an integral part a doctrine which has appeared since Hegel to be what Genova called a "philosophical liability": namely, Kant's transcendental idealism, and especially the reliance of this position on the noumenal/phenomenal distinction. Both Rawls and Habermas share as a common element to their respective meta-ethics an attempt to naturalize Kantian autonomy so as to rescue ethics from the need to posit the noumenal realm. Both these revisions make an attempt to tame the transcendental side of the conditions of autonomy; both Rawls and Habermas propose meta-ethical schemes that are ontologically much more austere than was Kant's. However, as I argued, both Rawls's and Habermas's meta-ethics leave unresolved questions whose answers must take recourse to the transcendental, however modest may be the contemporary interpretation of transcendentalism.

In the previous chapter I entertained what I take to be the most substantial and important criticisms of the central issue in the Kantian tradition of transcendental philosophy: transcendental arguments. I argued that, generally, while these critiques of TAs do apply to certain modern Kantian metaphysicians, the critiques generally misconstrue Kant. Most importantly, these critiques are generally inapplicable to the type of transcendentalism characteristic of the element of the neo-Kantian current in contemporary political theory that I

called ILT_p . In the terms of my general argument, in Chapter Three I argued that the critique of the Kantian account of autonomy that approaches autonomy from its transcendental side fails. The transcendental element of neo-Kantian political theory can withstand the contemporary attacks on transcendental arguments and the transcendental vindication of certain political values. In this chapter, I will entertain criticisms of neo-Kantian political theory that, roughly speaking, attack the neo-Kantian account of autonomy from the *political* side. I will argue that, in contrast with the critiques I took up in the previous chapter, these more political critiques of the neo-Kantian account of autonomy succeed in bringing into relief and challenging some fundamental assumptions of neo-Kantian political theory.

This distinction between the critiques of the philosophical and political elements of the neo-Kantian account of autonomy is, in a certain sense, imprecise. Rorty's critique of neo-Kantianism, which, in this context, I have treated as an attack on the 'transcendental' side of the neo-Kantian account of autonomy, certainly has political elements. A part of Rorty's critique of the transcendental vindication of political values is basically political; Rorty argues that invoking Kantian 'buttresses' impairs the process of developing the political virtue of solidarity based on the contingency of human social reality. However, I think that the critiques of neo-Kantian political philosophy offered by Foucault and some currents of

feminism share a certain 'empirical' element along with a kind of emphasis on the political consequences of neo-Kantian meta-ethics that is not shared by the types of critiques represented in Chapter Three by Stroud and Rorty. Foucault, in particular, has made it clear that his project was not predicated on a rejection of transcendentalism, but, rather, could be characterized, in this context, as an investigation into the consequences of methodologically suspending recourse to the transcendental.¹

The first two sections of this chapter will be given to an examination of a certain current of the feminist critique of neo-Kantianism, represented here primarily by Iris Marion Young, and the critique of neo-Kantian political philosophy that emerges out of Michel Foucault's writings, respectively. Both these perspectives attack the desirability of universality as a virtue of normative reason and as a mark of the moral. Young's and Foucault's arguments both criticize the political consequences of the adoption of universality as such a virtue and question the empirical viability of the assumptions such a meta-ethics makes

¹"In all of my work I strive...to avoid any reference to [the] transcendental as a condition of possibility for any knowledge. When I say I strive to avoid it, I don't mean that I am sure of succeeding. My procedure at this moment is of a regressive sort, I would say; I try to assume a greater and greater detachment in order to define the historical conditions and transformations of our knowledge. I try to historicize to the utmost in order to leave as little space as possible to the transcendental. I cannot exclude the possibility that one day I will have to confront an irreducible *residuum* which will be, in fact, the transcendental."

Foucault, Foucault Live, Sylvère Lotringer (ed.) (New York: Semiotext(e), 1989), p. 79.

concerning personhood. In particular, Young questions the assumptions underlying the adoption of universality as the mark of the moral, arguing that these assumptions inappropriately assign a morally arbitrary status to the body and physical and erotic desire. Young argues that the public sphere ought not always to have consensus-formation as its goal, but rather "the recognition and appreciation of differences, in the context of confrontation with power."² Foucault argues that conceiving of autonomy pre-politically--i.e., as a capacity to be preserved and respected in the political sphere--basically gets the problem backwards. 'Autonomy' is a politically informed value. Deeming a person or judgment 'autonomous' is a judgment generally informed by the extant theories of the human sciences. These theories, Foucault argues, cannot be fully understood if interpreted singularly as a collection of factual claims. The human sciences collectively produce a picture of normality: the healthy vs. the sick, the insane vs. the sane, the insane rather than the criminal, etc. By defining the normal, the human sciences define the permissible. Insofar as 'autonomous' action is action freed from pathological influences, the categories of the human sciences serve to delineate the conditions under which a subject's judgements may be considered autonomous, that is, free from pathological (e.g., criminal, mad, sexually perverse)

²Iris Marion Young, "Impartiality and the Civic Public: Some Implications of Feminist Critiques of Moral and Political Theory", Praxis International 5:4 (1986), p. 398.

influences. So, however different Foucault's analysis is from Young's, Foucault argues as well that securing the impartial realm of universalizable interests requires abstracting from the particularity or contingency of individuals in a politically normative and questionable way.

Both Young's and Foucault's critiques are directed at liberalism in general rather than neo-Kantianism in particular. In this context, 'liberalism' is both more broad and more narrow than 'neo-Kantianism.' It is more broad insofar as it 'liberalism' encompasses non-Kantian vindications of liberal values like tolerance and impartiality. It is more narrow insofar as 'neo-Kantianism' encompasses extra-political elements including epistemology and, in the form of ILT_p , a theory of truth. Nonetheless, Young's and Foucault's critiques of liberalism may be rendered as direct critiques of the neo-Kantian program as I have articulated and defended it. The first section of this chapter will be given to Young's critique of liberalism, the second to Foucault's critique, and the last to a consideration of the implications of Young's and Foucault's analyses on the question of the nature of autonomy.

Young's critique of liberalism

The type of critique of liberalism represented by Young's work differs in important ways from other, perhaps more popular forms of the feminist critique of liberalism. Young's critique, however, is the most telling and significant for the neo-Kantian

tradition. In hopes of bringing her arguments into sharper relief, I would like to introduce her work by way of contrast with two of these alternative feminist critiques. After this, I will turn to explicate Young's arguments in the context of a more detailed comparison with a third alternative feminist critique: Jennifer Nedelsky's programmatic recommendations for reconceiving the liberal concept of autonomy along feminist lines. Young's arguments are the focus of this review of various feminist critiques of liberalism; I bring up the alternatives for their utility in providing a background against which Young's arguments may be understood.

The first form of feminist critique I have in mind may be called a type of feminist naturalism. This critique is based on the type of challenge to normative rationality launched by the sociology of knowledge. The feminist form of the argument goes like this. What we call 'rationality' is just what we have taken, as a matter of fact, to be rational. 'Rational' is a compliment that has been paid to certain historically and politically contingent collections of practices and principles. As a matter of historical fact, "what we have taken to be rational" turns out to be "what *men* have taken to be rational." Also as a matter of historical fact, since the 17th century, the values of liberalism and certain core rational virtues have informed one another; e.g., 'impartiality' has been taken to be a necessary condition of a judgement's being rational, as well as a virtue of liberal jurisprudence. As well, central to the

epistemic and political elements of this tradition is a subordination of desire to reason. Insofar as women have been associated with desire and men with reason, the liberal tradition has been patriarchal.

As we will see, Young's argument has strong parallels with elements of this feminist naturalism, but it differs precisely on the grounds that Young's argument is not based on the naturalistic move in the first premise of the above argument (i.e., the reduction of de jure rationality to de facto rationality). She thus does not face the pitfall characteristic of the attempt to render compatible naturalism and normativity. The above argument must make such a move. The conclusion that, as a matter of fact, rationality and liberalism are patriarchal cannot, in itself, generate an imperative for reform. Such an argument can establish only that the distribution of power has been unequal, not inequitable.

The second type of feminist critique that I would like to contrast with Young's is a specifically feminist form of the communitarian critique of liberalism. The communitarian critique argues that the picture of subjectivity underlying liberalism is incorrect, and, once we abandon it, a new political ethics will emerge. This liberal theory of subjectivity is what Charles Taylor has called 'atomism': roughly, the picture of autonomous agents circumscribed by pre-political moral boundaries which imply 'given' rights which must be protected in political association. Communitarians argue that such a picture is simply

inaccurate, and call up, as an alternative, a neo-left-Hegelian account of subjectivity that argues that humans are, by nature, dependent on each other and on society for their identity.

'Autonomy' is a collective project. It is argued that from this it follows that we must reconstruct the polity to reflect the essentially social character of human subjectivity. This argument becomes a *feminist* argument with the addition of the thesis that the non-atomistic account of subjectivity³ is the specifically feminine (rather than gender-neutrally human) form of subjectivity. The establishment of a communitarian polity thus becomes a specifically feminist imperative.

Young's argument is directly opposed to this line. Her argument, roughly, is not that liberalism over-emphasizes autonomy, but that it understates it. Young argues that liberalism gains access to the abstracted autonomous subject that serves as its ethical core at the expense of abstracting from contingent persons a set of characteristics that is incorrectly taken to be morally arbitrary.⁴ Her critique is feminist insofar

³I choose "the non-atomistic account of subjectivity" rather than the more pithy "social account of subjectivity" because in feminist discourse, 'social' is often linked with 'public' and thus with the masculine side of the sexual division of labour in liberalism. According to this account, women are perhaps 'privately social'; forced into the private sphere politically, but more social by nature (i.e., by virtue of their role as child-rearers and care-givers). This distinction was pointed out to me by Thérèse Brabant.

⁴The point is a bit more subtle than this. As I have stated it, Young's misgivings are not with the process of abstraction, but with that which, in the liberal tradition, is abstracted. However--and in this sense, she bares a strong resemblance to Foucault--her critique is directed also at the *process* of

as she argues that, in the liberal tradition, the criteria by which characteristics are deemed morally arbitrary are informed by patriarchal values. I will examine her argument in some detail, considering also, along the way, Jennifer Nedelsky's programmatic suggestions for reconceiving autonomy along feminist lines. As we will see, Nedelsky's argument parallels and differs from Young's in a way that will both help broaden the base of Young's claims, and set them more clearly in relief against other feminist critiques of the theory and practice of liberalism.

Both Young and Nedelsky share a perspective on the values of liberalism and their relevance for contemporary politics that parallels to a certain degree Rorty's postmodern bourgeois liberal's agenda of disengaging the values of liberalism from the historical and theoretical context in which these values were first articulated and put into practice. Nedelsky argues that

Feminism requires a new conception of autonomy. The prevailing conception stands at the core of liberal theory and carries with it the individual characteristic of liberalism. Such a conception cannot meet the aspirations of feminist theory and is inconsistent with its methodology. The basic values of autonomy is, however, central to feminism. Feminist theory must retain the value, which rejecting its liberal incarnation.⁵

Similarly, Young argues that

There are plausible reasons for claiming that emancipatory politics should define itself as realizing

establishing personhood by methodologically abstracting the (supposedly) morally arbitrary aspects of contingent persons. More on this below.

⁵Jennifer Nedelsky, "Reconceiving Autonomy", Yale Journal of Law and Feminism 1:1 (1979), p. 7.

the potential of modern political ideals which have been suppressed by capitalism and bureaucratic institutions. No contemporary emancipatory politics wishes to reject the rule of law as opposed to whim or custom, or fails to embrace a commitment to preserving and deepening civil liberties...From the point of view of feminist interest, nevertheless, emancipatory politics entails a rejection of the modern tradition of moral and political life.⁶

Both Nedelsky and Young argue, that is, that the feminist critique of liberalism ought not to take the form of 'internal criticism.' The problem cannot be adequately addressed if we adopt as our critical framework the question of the extent to which liberal institutions have lived up to liberal ideals. That is, for different reasons, Nedelsky and Young argue that the task is not a matter of uniting liberal theory and liberal practice. Both Nedelsky and Young argue that when values like 'autonomy' are set within traditional liberal theory, their approximation in practice must involve practices of (unjust) exclusion.

Nedelsky's argument points to the historical association of autonomy with property. More than simply the fact that autonomy was associated with owning property, the autonomy-property link was the locus of the traditional liberal tension between the individual and the collective, and between democratic or procedural rights on the one hand, and substantive, pre-

⁶Iris Marion Young, "Impartiality and the Civic Public: Some Implications of Feminist Critiques of Moral and Political Theory", Praxis International 5:4 (1986), p. 381.

procedural rights on the other.⁷ In the 1780s, Nedelsky recounts, state legislatures started passing debtor relief laws "which were widely viewed as violations of property rights and as evidence of the intrinsic vulnerability of property (and, more generally, minority) rights under popular government." (Nedelsky, p. 16). In response, Federalists, drawing on Locke, sought to emphasize pre-politically guaranteed rights to autonomy (read: the ownership of property), arguing that these rights ought to be the object of legitimate government "and hence the limit to it." (Nedelsky, p. 17). This move, Nedelsky argues, solidified the conception of rights into the linked opposing categories of state vs. individual, public vs. private, and politics vs. the market. That is, 'autonomy' was cast as something which could be enjoyed at the point of a careful balance between the rights of the individual and the demands of the collective. 'Autonomy' thus became a private matter; the demands of the collective (and the results of democratic rule) were relegated to the public. And finally, the sphere of commerce in which autonomous action could reach articulation was the market: the free exchange of goods for capital. Politics, as opposed to the market, would be concerned with the proper administration of democratic rule and the preservation of the (private) right to autonomy.

⁷Historically, this latter tension can be explicated in terms of the difference between Hobbesian and Lockean contract theory. Roughly speaking, for Hobbes, the contract created individual rights that did not exist before agreement. For Locke, the contract functioned to protect pre-political or pre-consensual rights.

Nedelsky argues that, insofar as this conception of autonomy (and the linked dichotomies in which it reaches articulation and is set into political practice) is based on the link between autonomy and property, it must be abandoned. She gives two central reasons for the need for this revision. First, "the dichotomies of state-individual, public-private, politics-market, legislation-common law were always illusory." (Nedelsky, p. 18). Property was never an individual, private, market regulated institution in the first place. "[P]roperty rights are defined by the legal system. The security they provide rests on the power of the state to punish those who trespass on those rights." (Nedelsky, p. 18). Second, and more importantly, even if the autonomy-property link and the dichotomies that grow from it reflected some element of political reality in the past, property cannot play the role of locus of autonomy today. Property, Nedelsky argues, has lost its original political significance.

Property no longer provides people with the basis for independence and autonomy in the eighteenth-century sense. For the farmer who tilled his own land or the craftsman who owned his own tools, property was a real source of independence...[However] the dependence of wage earners on their employers is obvious. But even stockholders, who own their own shares, have little control over the source of their income. Their income, like that of most professionals, embeds them in a network of relationships characterized by interdependence rather than independence. (Nedelsky, p. 19).

So Nedelsky argues that we must abandon the model of property for autonomy, and with it, the strict distinctions between private-public, individual-state, and market-politics. Autonomy,

Nedelsky argues, must be conceived as a collective project, and not be relegated to the sphere circumscribed by the first half of these traditional liberal dichotomies.

If we ask ourselves what actually enables people to be autonomous, the answer is not isolation, but relationships--with parents, teachers, friends, loved ones--that provide the support and hindrance necessary for the development and experience of autonomy. I think, therefore, that the most promising model, symbol, or metaphor for autonomy is not property, but childrearing. (Nedelsky, p. 12.)

Young's critique of liberalism parallels Nedelsky's critique in aim and methodology. Like Nedelsky, Young wishes to disengage liberal values from their traditional theoretical and practical supports. And like Nedelsky--and unlike the 'communitarian' critique I reviewed above⁸--Young bases her critique on the political inadequacies of the classical liberal concept of

⁸Let me clarify the distinction I am drawing between Nedelsky's appeal for a non-individualist conception of autonomy and the communitarian critique of liberalism I reviewed above. The communitarian critique associated with Taylor, Sandel and Dallmayr is based on a more 'philosophical' concern than Nedelsky's argument for reconceiving autonomy. Taylor et.al. argue from a neo-left-Hegelian position that makes strong ontological commitments with respect to personhood, and argue, further, for a pseudo-deductive link between the theory of personhood and the theory of justice. Nedelsky's critique of liberalism does not draw on ontological misgivings concerning liberal individualism, but the discordance between the theory and practice of the liberal concept of autonomy. Again, however, her critique is not to be read as 'internal criticism'; she does not argue that we should reform liberal practice to live up to liberal theory. Rather, she argues, political reality demands that we reconceive autonomy along non-liberal lines.

Insofar as Nedelsky's argument is motivated by political concerns rather than misgivings about ontological assumptions, her critique has more in common methodologically with Young's than with the communitarian critique associated with the neo-left-Hegelians.

autonomy. Nedelsky and Young agree, further, that the traditional public/private distinction is inaccurate, that the so-called 'private' sphere has always been political, and that the first step in reconceiving autonomy is realizing the intrusion of politics into the political sphere. However, for Young, the fact that the private sphere is informed by politics (pace the traditional liberal alignment of market-politics with private-public) does not imply that autonomy be reconceived along communitarian lines.

In sloganistic terms, Young argues that feminists (and radical political theorists and activists in general) should abandon the 'ideal of community' and embrace a 'politics of difference.' In a nutshell, her argument goes as follows. The idea of a community requires a certain abstraction from the contingencies of its members. Replacing the individualist social ontology of liberalism with a more communitarian conception of personhood and personal interaction only perpetuates the underside of the liberal public/private distinction. By relegating certain elements of personhood--those, it turns out, most associated with developing autonomy--to the private realm, liberal politics also demands that the concerns related to these private elements of personhood remain private; remain, that is, excluded from the public and political sphere. The problem with the liberal conception of autonomy, Young argues, is not that it over-emphasizes individualism, but that the 'individuals' over which liberal theory quantifies are contrived entities, arrived

at by a process of abstracting the elements of personhood from contingent individuals that are deemed morally arbitrary. The dark side of liberal politics is that the establishment of a public realm consisting of individuals thus construed is predicated on the exclusion of morally and politically relevant elements of personhood.

Young's argument is that this exclusion underscoring the liberal polity is not accidental. It is not a matter, that is, of creating more space in which the voices of the excluded can be heard. The liberal concept of autonomy and personhood is predicated on this procedure of exclusion. This follows, Young argues, from the fact that arriving at the concept of personhood underscoring the liberal individual requires setting into opposition elements of personhood: essence-accident, normal-deviant, and, most importantly, mind-body. The problem with this is that

The dichotomies are not symmetrical, however, but stand in hierarchy; the first term designates the positive unity on the inside, the second less valued term designates the left-over outside. (Young, p. 386)

The liberal value of equality, for example, prides itself on an indifference with respect to bodily qualities; skin colour, appearance and gender are irrelevant attributes with respect to the rights of citizenship. The problem, Young argues, is that the flip-side of this positive claim is the effective exclusion of 'body-issues', e.g., homosexual rights, racial minority rights, from public debate. Insofar as women have been

associated with the body--as child-bearers, as house-keepers--
and

desire, issues central to the liberation of women (along with other marginalized groups) have effectively been excluded from public, and hence political representation.

Like Nedelsky, Young is not attempting to overturn the value of privacy. Unlike Nedelsky, however, Young does not conclude that the solution to the hidden inequalities of the liberal conception of autonomy lies in recasting autonomy in communitarian terms.

Instead of defining privacy as what the public excludes, privacy should be defined, as an aspect of liberal theory does, as that aspect of his or her life that any individual has a right to exclude from others. I mean here to emphasize the direction of agency, as the individual withdrawing rather than being kept out. (Young, p. 396).

So the intrusion of politics into private life does not imply for Young that we should dissolve the public/private distinction, as communitarian critiques of liberalism imply. Rather,

The feminist slogan, "the personal is political" does not deny a distinction between public and private, but it does deny a social division between public and private spheres, with different kinds of institutions, activities, and human attributes. Two principles follow from this slogan: (a) no social institutions or practices should be excluded a priori as being the proper subject for public discussion and expression, and (b) no persons, actions or aspects of a person's life should be forced into privacy. (Young, p. 396).

How does Young's critique apply to the neo-Kantians I have thus far been defending? I would suggest that, most significantly, Young's critique of liberalism draws into question

both the idea of universality as a mark of the moral and the viability of 'rational consensus' as a political goal, or as an ideal towards which the just polity ought to aspire. Recall the basic elements of Kantian meta-ethics that remained at the core of contemporary neo-Kantian political theory. For a judgment to be binding, and thus moral, it must be universalizable. It must be free, that is, from pathological influences. Now, the 'pathological' has undergone a great deal of revision since Kant. In the current of neo-Kantianism I have been defending--ILT_p--the 'pathological' is interpreted in purely procedural terms, basically as constraint on the process of consensus-formation. So while Kantian meta-ethics undergoes substantial revision in the hands of neo-Kantian ideal limit theorists, the idea of 'universalization' as the mark of the moral is unchallenged, but simply recast in terms of 'ideal' or 'rational' consensus.

Young's argument suggests that it is *precisely* that which does not admit of universalization--what is, in fact, in neo-Kantian terms, pathological--that should be the object of our moral concerns and political agendas. Her argument suggests, that is, that it is exactly those interests that are tied to the contingent, non-universalizable aspects of personhood--most importantly, those aspects of personhood associated with the body, and thus central to the concerns of, e.g., visible minorities, women, homosexuals--that must be addressed in a just society. In these terms, according to Young's critique, ILT_p has made significant progress in the procedural revision of Kantian

meta-ethics, insofar as Habermas et al. make interests the objects of consensus-formation. (Recall that for Kant the pathological was basically material interest). However, according to Young, ILT_p simply re-produces the classical liberal distinction between public and private (and, along with it, other dichotomies central to liberalism, most importantly desire and reason). An interest which is non-universalizable, even if it be excluded qua non-universalizable (Habermas), rather than qua interest (Kant), is deemed by ILT_p to be inappropriate for public and thus political concern. In Young's terms, ILT_p maintains the liberal allocation of agency with respect to the question of determining what is to count as private in the hands of the collective decision-making process, rather than placing such agency in the hands of the individual. The two principles that Young argues follow from the feminist slogan "the private is political"--that (a) no social institutions or practices should be excluded a priori as being the proper subject for public discussion and expression, and (b) no persons, actions or aspects of a person's life should be forced into privacy--are violated by ILT_p .

I think that Young has pointed to a significant political inadequacy in the neo-Kantian program, and, furthermore, has brought into relief a strong traditional element in neo-Kantian politics. Before assessing the implications of her critique for ILT_p , however, I will turn to review a second powerful critique

of neo-Kantianism, found in the works of Michel Foucault. Drawing on very different material, Foucault's critique of liberalism similarly points to a hidden system of exclusion underscoring the positive liberal account of autonomy.

Foucault's critique of liberalism

Foucault's critique of the liberal conception of rights and equality points to a hidden system of inequality that is intertwined with the formal guarantees of equality. His critique is rooted in the claim that the liberal conception of rights misrepresents the nature of power by thinking of power exhaustively in "economistic" terms⁹: as something to which one has a right which can be possessed like a commodity, and thus transferred like a commodity¹⁰; or as the "privilege", acquired or preserved, of the dominant class¹¹; or as something which can be "checked" or "limited" by the establishment of rights by the law (PK, pp.105-107); or as something which serves as a limit set

⁹not "economic" terms, in the sense of the political/economic division central to liberal capitalism.

¹⁰Michel Foucault, Power/Knowledge: Selected Interview and Other Writings, Colin Gordon (ed.), (New York: Pantheon Books, 1980), p. 88. All further references to this text will be inserted into my text in the form: (PK, p.__).

¹¹Michel Foucault, Discipline and Punish: The Birth of the Prison (New York: Vintage Books, 1979), p. 26. All further references to this text will be inserted into my text in the form: (DP, p.__).

on freedom¹²; or as something exclusively found in the state (PK, p. 39, 122).

This "economistic" understanding of power, Foucault argues, underscores the "juridical" account of power. By the "juridical" account of power, Foucault means that account that understands power in terms of rights--one may have the right to exercise power, and a right may check the exercise of power. Power (and rights) may be transferred (in a contract) or alienated (with the creation of a sovereign). This is, essentially, the classical liberal theory of right. Foucault's claim is that this account is incomplete, that "we should not be deceived by all the Constitutions framed throughout the world since the French Revolution, the Codes written and revised, a whole continual and clamorous activity: these were the forms that made an essentially normalizing power acceptable (HS, p. 144)." This "essentially normalizing power" is the power of the disciplines, the "techniques for assuring the ordering of human multiplicities (DP, p. 218)." These "techniques" are the techniques of the human sciences (psychology, psychiatry, pedagogy, criminology, anthropology, medicine, sociology, political science, economics). They determine, essentially, what is required to *qualify* as a subject competent to enjoy these rights, and, as such, complement the formal requirement of homogeneity demanded by the liberal

¹²Michel Foucault, The History of Sexuality (vol. 1) (New York: Vintage Books, 1980), p. 86. All further references to this text will be inserted into my text in the form: (HS, p. __).

theory of rights by supplying a practical method of individuation supported by the findings of the human sciences (PK, pp. 106-107).

The "Enlightenment", which discovered the liberties, also invented the disciplines...[W]hereas the juridical systems define juridical subjects according to universal norms, the disciplines characterize, classify, specialize; they distribute along a scale, around a norm, heirarchize individuals in relations to one another and, if necessary, disqualify and invalidate. (DP, p. 223)

The system of right has been superimposed upon the mechanism of disciplinary power in such a way "as to conceal its actual procedures, the element of domination inherent in its techniques, and to guarantee to everyone, by virtue of the sovereignty of the State, the exercise of his proper sovereign rights (PK, p.105)." The acceptability of power lies in its ability to mask a substantial portion of itself (HS, p. 86).

Foucault emphasizes that his account of power is not reducible to an analysis that casts power in terms of the hegemony of the ruling class. Foucault's critique of posing the problem in terms of class analysis is Popper-like¹³: the problem lies not in a lack of explanatory power, but in its excess. One can always explain phenomena away by 'deducing' the explanation from the dominance of the bourgeois class--the fact that the bourgeois is dominant guarantees the 'correctness' of all such deductions (PK, p.100). The problem, rather, is to see how

¹³cf. Karl Popper, "Science: Conjectures and Refutations", Janet A. Kourany (ed.), Scientific Knowledge (Belmont: Wadsworth Publishing Co., 1987), pp. 139-157.

"mechanisms of power, at a given moment, in a precise conjecture and by means of a certain number of transformations, have become economically advantageous and politically useful (PK, p. 101)."

These mechanisms of power are inseparable from claims to knowledge. What precisely is the relationship between power and knowledge? There are, as Ian Hacking suggests, two bad short answers (neither of which is Foucault's, but are rather, Hacking suggests, misconstruals of Foucault): "(1) knowledge provides an instrument that those in power can wield for their own ends; (2) a new body of knowledge brings into being a new class of people or institutions that can exercise a new kind of power."¹⁴ We may be tempted to assert either of these equations as an explanation of the maintenance or emergence of the bourgeoisie as the ruling class by linking this emergence to the possession of a new science of economics that made possible the accumulation of capital (which in turn secured their dominance as owners of the means of production). Foucault suggests, however, that the process of the accumulation of capital--made possible by the development of a new mode of production--cannot be separated from the process of the "accumulation of men" (DP p. 221), which was made possible by disciplinary power. The practical problems involved in making the most efficient use of the new commodity of labour power and organizing the division of labour--the design of

¹⁴Ian Hacking, "The Archaeology of Foucault", David Couzens Hoy (ed.), Foucault: A Critical Reader (London: Basil Blackwell, 1986), p.27.

factories, the introduction of the work-day--were solved by "a 'knowledge' of the body that is not exactly the science of its functioning, and a mastery of its forces that is more than the ability to conquer them (DP, p. 26)," that "permits time and labour, rather than wealth and commodities, to be extracted from bodies (PK, p. 104)."

Foucault's point is that this ambiguous power/knowledge relationship¹⁵ permeates our institutions: hospitals, schools, places of work, prisons, courts. The classifications and norms that serve on the one hand to divide e.g., the healthy from the sick, the mentally from the physically ill, the criminal from the insane, the productive from the unproductive, and on the other, to regulate, observe, and discipline the patients, students, workers and criminals seem to be simultaneously claims to truth by the human sciences and techniques of power exercised by society. How does this differ from both of Hacking's "bad short answers"? I would suggest this is illustrated by Hacking's claim that both formulations--(1) knowledge provides an instrument that those in power can wield for their own ends; and (2) a new body of knowledge brings into being a new class of people or institutions that can exercise a new kind of power--parallel two opposed theses about ideology:

- (1) a ruling class generates an ideology that suits its own interests; and (2) a new ideology, with new values,

¹⁵i.e., the one that characterizes "a 'knowledge' of the body that is not exactly the science of its functioning, and a mastery of its forces that is more than the ability to conquer them"

creates a niche for a new ruling class. (Hacking, p. 27)

But Foucault's point is that the question of the relationship between power and knowledge is not a matter of ideology.

[I]t is quite possible that the major mechanisms of power have been accompanied by ideological production...but I do not believe that what has taken place can be said to be ideological. It is both more and much less than ideology. It is the production of effective instruments for the formation and accumulation of knowledge--methods of observation, techniques of registration, procedures for investigation and research, apparatuses of control. All this means that power, when it is exercised through these subtle mechanisms, cannot but evolve, organise and put into circulation a knowledge, or rather apparatuses of knowledge, which are not ideological constructs. (PK, p. 102)

Thus, Foucault suggests, "the political problem, to sum up, is not error, illusion, alienated consciousness or ideology; it is truth itself (PK, p.133)." The problem does not consist "in drawing the line between that in a discourse which falls under the category of scientificity or truth, and that which comes under some other category, but in seeing how effects of truth are produced within discourses which in themselves are neither true nor false (PK, p. 118)."

Let me spell out these last few claims. The concept of ideology against which Foucault contrasts his power/knowledge

thesis is neither the orthodox Marxist conception that I reviewed in Chapter Two, nor Habermas's revision of this concept in light of advanced capitalism. The interpretation of 'ideology' against which Foucault sets his views is the more popular received notion of ideology as a false world-view perpetuated by the ruling class in the interests of maintaining the extant mode of production. Given this interpretation of 'ideology,' Foucault claims that the operation of disciplinary power is *more* than ideology because these operations are not simply misrepresentations of material interests, or institutions put into place to defend a certain political order; rather, they make possible the very functioning of a society not ruled by the (explicit) sword. At the same time, the functioning of disciplinary power is *less* than ideology because the knowledge claims of the disciplines do not provide a coherent world view whereby a given order is legitimized; they are scattered, discontinuous, often isolated. The "knowledge" produced by the disciplines is not *illusory* as ideological knowledge is; it is not power hiding under the cloak of knowledge. The "production of effects of truth" is the realization or materialization of the truth claims of the disciplines in practice; the disciplines *discipline*. By stating that this "production" occurs within "discourses which are themselves neither true nor false", Foucault is disallowing an appeal outside of the discourses of the disciplines that would wish to adjudicate over the truth and falsity of their claims. The 'truth' of these claims is determined by their realization in

practice. That is, Foucault is arguing that there is no formal test by which a given claim in the human sciences is accepted as 'true', in the sense that, e.g., ILT_p seems to suggest. Rather, Foucault would argue, talk of 'assigning the t-value "true" to a proposition p ' does not adequately describe the process of a proposition's coming to be regarded as 'true' within the human sciences. The criterion by which a given proposition may be regarded as true within the human sciences--if it is even appropriate to read something as formal as a 'criterion' out of Foucault's account of power/knowledge--is more like: ' p is true iff the normative consequences of adopting p (i.e., the practical implications for the institutional practice of human science to the extent that regarding p as true has regulative consequences) accords with the extant structures of power in society (not 'structures of power' in the sense of the ruling class, but in the sense in which individuals are classified, and, on the basis of this classification, regulated).'¹⁶

¹⁶It should be noted that Foucault's analysis does not exclude the possibility that scientific theories may nevertheless be true, in the sense of 'truth-as-correspondence.' This point is analogous to the point I made in Chapter Three that any epistemic theory of truth is logically compatible with metaphysical realism: i.e., whatever we may say about our epistemic practices and our practice of assigning truth-value 't' to a given proposition, it is always possible that our extant theories match up with a mind-and discourse-independent reality to a greater or lesser degree. Similarly, Foucault's analysis does not rule out the possibility that a similar 'mapping' occurs with the human sciences to a greater or lesser degree.

Foucault himself also noted that his analysis is not aimed at rejecting the claim that the human sciences have not satisfied some locally specified practical ends:

"My problem is...to know how games of truth can put themselves in place and be linked to relationships of power. We

Now, unlike (at least certain popular currents of) Marxism, Foucault is not attempting to contrast, say, claims advanced in the interests of preserving power vs. claims advanced in the service of furthering the pursuit of truth. Foucault's claim about the power/knowledge relationship in the disciplines is an explicit rejection of the "science on the one hand/the ideological use of science on the other" disjunction. This is where the distance of his analysis from a critique of ideology is most clear and relevant. He is not making any claims about who the disciplines serve, or that they serve anybody in particular; the "power/knowledge" thesis does not reduce science to one grand Lysenko affair. It is not that power cloaks itself with claims to "knowledge" as a means of legitimation; Foucault's analysis is not aimed at disclosing an 'occult Machiavellianism'¹⁷ at work in the human sciences. His claim is, rather, that the very establishment of claims as "true" (and thus legitimate) is simultaneously an exercise of power--a power exercised, not by those who put forward these claims, but by the specific content

can show, for example, that the medicalization of madness, i.e., the organization of medical knowledge around individuals labelled as "mad," has been linked, at some time or other, to a whole series of social or economic processes, but also institutions and practices of power. This fact in no way impairs the scientific validity of the therapeutic efficacy of psychiatry. It does not guarantee it but it does not cancel it out either."

Foucault, "The Ethic of Care for the Self as a Practice of Freedom", James Bernauer and David Rasmussen (eds.), The Final Foucault (Cambridge: MIT Press, 1988), p. 16.

¹⁷A phrase used by Foulek Ringelheim in an interview with Foucault. "What Calls for Punishment?", reprinted in Foucault Live, p. 284.

of these claims themselves (whatever they may be at a particular time): by defining the normal, the disciplines define the permissible.

I would argue that, in a manner similar to Young's critique of liberalism, Foucault's analysis of the power/knowledge relationship at work in the human sciences and the role the human sciences play in our received notions of personhood draws into question the neo-Kantian exclusion of non-universalizable interests from matters of justice. Again like Young, Foucault suggests to us that the individuals over which liberalism quantifies when it speaks of equality of persons before the law are delimited by means of a process of exclusion. Unlike Young, Foucault points not to the exclusion of the body at the theoretical level, but at the exclusion of particular person at the political level, e.g, the insane, the sick, the criminal. The theory of right, Foucault argues, is superimposed upon the practice of disciplines.

I would argue that it follows from Foucault's analysis that autonomy is not simply a political right, nor is it something that we can speak of as guaranteed by human nature, or secured by the structures of human commerce. To act autonomously is, among other things, to be able to take responsibility for one's actions. It is also, however, to be allowed to take responsibility for one's actions. As Young's analysis of liberalism pointed to a misplacement of the position of agency with respect to the decision of what aspects of a person's life

are to remain private, Foucault's analysis seems to point to a certain misplacement as to the locus of agency with respect to the decision of assuming responsibility for one's actions. That is, e.g., the mad are not simply *forgiven* the burden of assuming responsibility for themselves, they are *forbidden* to assume such responsibility. The human sciences' tool kit carries a double-edged sword.

Foucault's analysis thus leads to questioning the neo-Kantian adoption of Kant's formula of the moral judgement as the judgement free from pathology. Even when 'pathological judgement' becomes reinterpreted along purely procedural lines, as with Habermas, the idea of a non-pathological discursive situation (the 'ideal speech situation') is predicated (in Habermas's analysis) on the idea of a collection of individuals who may be quantified over as equals (in this case, equal participants in the dialogue, rather than equal bearers of rights). Foucault's analysis grounds Young's more theoretical misgivings about the process of abstraction upon which such a quantification must be based in a more empirical study of the means by which we--in the age in the human sciences--classify and order subjects. Foucault's analysis, in short, undermines any pretensions to a purely descriptive (or politically inconsequentially normative) procedure whereby persons are granted the competent status required for participation in debate in the public realm. In Habermas's terms, I would suggest that Foucault might argue that procedural pathology cuts deeper than

Habermas's analysis points to. The process of exclusion, that is, occurs not only at the level at which persons may assert themselves in public, but at the level at which persons are deemed qualified to enter into such participation. As with Young's critique of liberalism, I would argue that it follows from Foucault's analysis of the disciplines that the equation of the moral with universalizable interests may be approximated in practice in the form of public consensus-building only if the mechanisms that secure persons 'equal' participation in the dialogue are superimposed upon a system of inequity and exclusion. It follows from this, as with Young's analysis, that, *pace* ILT_p, it is perhaps precisely the interests of persons that are non-universalizable that ought to be the objects of our moral concern and agendas for political reform.

Conclusion: The nature of autonomy

The Young-Foucault critique of neo-Kantian political theory¹⁸ brings to the fore the question of whether autonomy ought to be cast in procedural or in substantive terms, or as some balance between the two. As Nedelsky points out, the idea

¹⁸Let me qualify my use of the phrase 'the Young-Foucault critique'. I am not suggesting that Young and Foucault were allied, nor even that they agree on precisely what the problems are in the neo-Kantian approach. But, as I argued above, I think that together, their analyses offer a broadly-based critique of the neo-Kantians. So please read 'the Young-Foucault critique' as a phrase capturing what I have argued to be the complimentary elements of their respective critiques of liberalism and neo-Kantianism.

of autonomy as a balance between procedural and substantive elements is central to liberal theory (Nedelsky, pp. 16-18). That is, while acting autonomously involves full participation in collective decision-making, at the same time, it seems to require, as well, the protection of certain rights against what Mill called the "tyranny of the majority."¹⁹ Young and Foucault seem to invite a contemporary version of the fear Nedelsky tells us was felt by American Federalists at the turn of the century concerning popular tyranny: that the procedural mechanisms of representation allow structural space wherein the claims of the individual may be overruled by the claims of the collective. The contemporary version of this fear does not derive from the possibility allowed by the democratic process of the emergence of a *quantitative* collective that threatens the voice of the individual. It derives, rather, from the inability of the formal mechanisms of political representation--however carefully constructed--to prevent what we might call a *qualitatively* exclusive collective body from exercising political hegemony. That is, the Young-Foucault critique of a purely procedural interpretation of autonomy does not attack the notion of pure procedural justice on behalf of the quantitative minority who lose in the process of majority rule. Their concern, rather, is with, on the one hand, those excluded from the process--those whose interests concern the body (Young); those who fail the

¹⁹John Stuart Mill, On Liberty (Illinois: Harlan Davidson, 1947), p. 4.

competency tests of the human sciences (Foucault)--, and, on the other hand, the mechanisms that must be in place to secure the homogeneity required among the dialogue-participants so that they may enjoy the equitable distribution of rights--the systematic exclusion of the body and its concerns (Young); the policing of behaviour (Foucault). In short, I am arguing that we can read Young and Foucault not as concerned to question the possibility of an inequitable or unjust procedural *outcome* (as are liberal theorists who emphasize pre-political rights), but as concerned rather to uncover the theoretical and practical mechanisms required to secure a body of citizens who are 'equal before the law', or are 'equal participants in consensus-formation.'

Two significant questions may be raised about Young and Foucault, neither of which is easily answered. (1) Insofar as Young and Foucault challenge the procedural interpretation of autonomy, are they more than present-day analogues of the American Federalists? That is, do their analyses not simply point to the need to revise the procedural-substantive balance in the liberal conception of autonomy (and thus, are their criticisms merely internal to liberalism)? (2) If the traditional elements of the substantive side of the liberal conception of autonomy (e.g., pre-political rights à la Locke) are predicated on a process of exclusion, or are merely superimposed upon the mechanisms of discipline, but a purely procedural conception of autonomy is inadequate, then what substantive assumptions must play the role that is played by

traditional liberal substantive assumptions in the concept of autonomy? In short, if the liberal tradition was incorrect about the substantive assumptions admissible within a theory of autonomy, but correct about the need to include *some* such assumptions, then what ought these assumptions to be?

These questions cannot be clearly disengaged from each other. In simplest form their answers are : (1) Young and Foucault are not internal critics of liberalism, because (2) while they point to the inadequacy of a procedural conception of autonomy, their critiques also undermine the idea of pre-procedurally established substantive claims concerning personhood and autonomy. Let me try to spell this out.

To read Young and Foucault as internal critics of liberalism is to miss the point of their critiques. Young and Foucault are not pointing to ways in which the practice of liberal politics has not lived up to the ideals of liberal theory. They are arguing, rather, that the ideals of liberal theory demand an unrecognized system of exclusion to allow these ideals to be approximated in practice. To the extent that such a process of exclusion is incompatible with liberal values, then we can read Young and Foucault as internal critics of liberalism. Their point, however, is that these processes of exclusion are necessitated by the process of realizing liberal ideas in practice.

The second question above seems to come down to the issue of whether Young and Foucault are making categorical claims about

the process of defending pre-procedurally articulated rights. That is, does it follow from the Young-Foucault critique that any attempt at universalizing interest must involve a process of exclusion embedded in a system of normalizing power? Or is it, rather, that once we abandon the liberal framework, and with it the assumptions concerning personhood that inform the liberal conception of autonomy, we will be able to articulate a conception of autonomy that is free from the need to invest itself in mechanisms of exclusion? At this point, Young's and Foucault's analyses are merely suggestive. While, on the one hand, it seems unjustifiably visionary to suppose that some post-liberal politics will overcome the liberal investment in processes of exclusion, on the other hand it seems that arguing that no possible political organization conceivable can avoid these trappings precisely involves what Rorty would call an unjustifiably generalized parasitism argument. That is, a categorical claim to the effect that, in Foucault's terms, all political organization must invest in disciplinary power (or some analogue), seems to stake a claim to a transcendental status that would be difficult to defend.

Both of the questions I asked of Young and Foucault above assume that the procedural/substantive distinction is exhaustive with respect to conceptions of autonomy. Further, these questions, as well as my discussion in general, also assume that the procedural/substantive distinction may be drawn, if not precisely, then adequately. Perhaps the confusion and difficulty

surrounding the attempt to make Young's and Foucault's analyses answer these two questions is occasioned by the fact that their analyses point to the procedural/substantive distinction as a liberal dichotomy to be challenged along with the challenge to public/private, politics/market, etc. This, I think, is an enormously suggestive and perplexing outcome. What, exactly, would it mean to think of autonomy in non-procedural or non-substantive terms? What would it mean, that is, to talk about autonomy and make no reference either to the ability of a subject to engage in decision-making, or to the ability of a subject to enjoy certain rights with respect to the demands of others? The problem here, it seems, is that we cannot formulate the idea of autonomy without first assuming some level of personhood, at least to the point that 'persons' may be quantified over (or, if the concept of 'quantification' is troublesome, or carries with it philosophical concerns extraneous to this discussion, then substitute: at least to the point that 'persons' has a place in our language-game). Young and Foucault seem to be attacking this assumption, or the ease with which we speak of persons who are or are not autonomous. The Young-Foucault critique suggests that there is, in liberalism, a political investment in the issue of personhood, and that the question of 'autonomy' is at least dependent on the question of what or who is to count, in political terms, as a person (citizen, subject, etc.). Their critique draws into question the procedural/substantive distinction not so much at the level of the distinction as it

applies to persons, but, rather, insofar as the distinction depends upon a non-controversial concept of personhood. Young and Foucault both question an assumption common to traditional liberal and contemporary neo-Kantian political theory: that politics intrudes into the issue of autonomy *after* the establishment of the nature of personhood. This assumption is left undefended by neo-Kantian political theory. I would argue that the positive program of political critique defended by neo-Kantianism must come to terms with this challenge.

In sum, then, I have argued that neo-Kantian political theory has successfully naturalized Kant's account of autonomy, but has done so only with 'transcendental residue.' Neo-Kantians share with Kant the idea that autonomy has both transcendental and political elements, neither of which can be reduced to the other. I have argued that the attack on the neo-Kantian account of autonomy from the transcendental side fails to the extent that the critiques of transcendental arguments and transcendental vindications of political values in current literature do not apply to the modest transcendentalism that characterized ILT_p. However, I have argued, there is a powerful critique to the political side of the neo-Kantian account of autonomy that has not yet been but which must be answered by neo-Kantianism, if the contemporary vindication of the critique of ideology is to be defensible.

CONCLUSION

This essay is basically a review of some of the major currents in contemporary political theory. It has been my primary objective to argue that we can find a certain unity among many of the apparently disparate positions in contemporary political theory if we collect these positions under the rubric of certain questions left to us by Kant. I think that such a project of historical reconstruction is useful because, as Charles Taylor has argued, we are afforded greater understanding of received ideas when we recover their original formulations. In this context, given that I review French and German as well as Anglo-American theorists, I also offer such a reconstruction in an attempt to further the bridging of the analytic-continental gap in contemporary philosophy.

In its most basic form, the Kantian picture that I have argued still has a hold on contemporary political theory can be expressed formulaically thus:

§1 Certain philosophical conditions + certain political conditions = autonomy;

§2 Autonomy = the grounding of political critique.

The bulk of my thesis has been an examination of the debates surrounding §1: what the philosophical and political conditions that serve as preconditions for autonomy must be (Rawls vs. Habermas); whether these philosophical conditions can be transcendentally vindicated (Stroud and Rorty vs. Habermas); whether we need to appeal to these philosophical conditions

(Rorty vs. ILT_p); whether we can formally delimit these political conditions (Young and Foucault vs. ILT_p). It has not been a matter of editorial decision to restrict my discussion of the current literature to an examination of the various issues surrounding §1, to the exclusion of §2. It is rather, I would argue, a reflection of the current state of political theory. The significance of Kant for contemporary political theory is perhaps best illustrated by the fact that the link between autonomy and the grounding of political critique (§2) has remained, by and large, unchallenged. I think that Young's and Foucault's critiques, while perhaps more directly addressing the formula given in §1 than in §2, also point to the beginnings of such a challenge to the second part of the Kantian formula. I think, further, that this challenge to the autonomy-political critique link is the most important issue emerging out of contemporary political theory. But before I spell out what I think this issue amounts to (that is, the issues surrounding §2), let me review the argument in the chapters above (that is, the debates surrounding §1).

In Chapter One, drawing on O'Neill and--to a lesser degree--Foucault, I defended a reading of Kant that proposes that we not look to Kant's writings to find a political theory, but rather that we recognize a political element to the whole critical philosophy. The heart of Kant's critique of pure reason is the Copernican principle: the claim that knowledge is restricted to possible experience, and that we can learn from experience only

as much as the structure of reason allows us. In contemporary terms, Kant was proposing (and was perhaps, as Putnam suggests, the first to propose) a purely epistemological theory of truth. Kant thus placed certain traditional metaphysical concerns outside the boundaries of legitimate knowledge, insofar as these concerns could not be settled by an appeal to possible experience. Among the issues from which knowledge was thus prohibited by Kant was the question of the extent to which our actions as moral agents is free from determination by causes outside ourselves. Kant argued that we could never know--in the only sense that 'knowing' is possible--whether we were free in this sense, because the issue could never be settled by appeal to experience. However, Kant gave practical reason (ethics) the licence he had revoked from theoretical reason (epistemology). Insofar as ethics is concerned with what ought to be, rather than what is, ethics is not constrained to operate within the boundaries of the empirically verifiable. As such, in ethics, we are justified to assume that our actions are free--and thus, we are bound to assume responsibility for our actions. However, to be truly autonomous, Kant argued, we must not only possess free will, but the free will must have political space in which to act. Autonomy, for Kant, thus has two components; it is both a transcendental and a political matter.

In contemporary terms, we can say that, according to Kant, if we assume an epistemological theory of truth--that is, if we regard 'truth' as a function of the satisfaction our epistemic

criteria, rather than a function of the extent to which our epistemic activities map the world which we seek to describe-- then we are led to regard as the final tribunal of our epistemic claims and moral judgements the rational consensus of our fellow inquirers. Now, for Kant, the 'rational' component of 'rational consensus' was defined in accordance with the idea of judgement absolutely free from material interest. Freedom from material interest was conceivable, according to Kant, only inasmuch as we could conceive of ourselves as noumenal beings--as things-in-themselves independent of time and space. Contemporary neo-Kantian political philosophy has at its core an attempt to revise Kant so as to keep the basic structure of his meta-ethics, but do away with the equation of rational judgement with freedom from material interests and residence in the noumenal realm. Rawls and Habermas propose two distinct accounts of rational judgement that serve the role played in Kant's meta-ethics by the noumenal realm, but do so with less ontological extravagance: the veil of ignorance, and the ideal speech situation. For both Rawls and Habermas, as well, rational judgement no longer appears as freedom from material interest, but freedom from ill-defined material interest. The veil of ignorance and the ideal speech situation serve, for Rawls and Habermas, respectively, as regulative ideals that serve to counter the pathological influences on rational judgement. Rawls reinterprets the pathological substantively, identifying it with the intrusion of the morally arbitrary into moral judgement. Habermas

reinterprets the pathological procedurally, identifying it with structural constraints on the process of consensus-formation.

Neither Rawls nor Habermas, however, can complete their naturalization of Kantian meta-ethics without transcendental residue. That is, while Rawls and Habermas argue ably for the moral implications of impartial reasoning, neither Rawls nor Habermas can explain that we ought to be rational. As Habermas recognizes explicitly, and Rawls implicitly, such an explanation must take recourse to the transcendental. That is, both Rawls and Habermas are constrained to argue that moral agents do not have the option to 'opt out' of behaving rationally; they must argue, that is, that acting at all commits one to the schemes of rationality they propose. Now, the idea of transcendental philosophy has come under sustained attack from various positions in contemporary philosophy. I argued that the kind of transcendental argument required by contemporary neo-Kantian political philosophy--distinguished from traditional transcendental arguments insofar as the position to be defeated by these contemporary transcendentalists is not the epistemic sceptic, but rather the moral nihilist--successfully escapes the major criticisms in the literature. I argued, further, that this form of neo-Kantian political philosophy can be described as a procedural interpretation of ideal limit theory. By this I mean to say that at the core of the neo-Kantian tradition represented by Habermas, Apel, and Putnam was the idea that our epistemic and moral judgments had as their final tribunal the judgements of our

fellow-inquirers under conditions wherein consensus-formation was free from coercion. I argued, finally, that, while Rorty's misgivings about neo-Kantianism give us good reasons to prefer ILT_p over the Rawlsian revision of Kant, Rorty's misgivings concerning ideal limit theory did not apply to the procedural interpretation proposed by Habermas and company. I suggested, finally, that Rorty's anti-transcendental pragmatism is not nearly as distant from ILT_p on the philosophical map as Rorty would like to suppose, and, further, that Rorty's 'compliment' theory of truth, *pace* Rorty, commits us to the kind of ideal limit that Habermas, Apel, and Putnam ask us to accept.

In terms of my general argument, I defended neo-Kantian political philosophy against the critics who attack the neo-Kantian account of autonomy from the transcendental side. However, in the final chapter, I argued that neo-Kantianism is not so immune from the criticisms launched by certain theorists against the political side of the ILT_p account of autonomy. Young and Foucault argue that 'rational consensus' as a standard applied to our epistemic and moral practices can inform the construction of the polity only at the expense of a political investment in mechanisms of exclusion. Young argues that the equality required among participants in consensus-formation can be achieved only by excluding morally relevant elements of personhood. Foucault argues that particular persons must be excluded to produce the homogeneity required of a populous over which rights may be distributed equally. In short, Young and

Foucault argue that politics intervenes not only at the level of autonomy, as the neo-Kantian recognize, but at the prior level of personhood.

I would argue that Young's and Foucault's critiques of neo-Kantianism address not only the neo-Kantian account of autonomy (i.e., §1 in the formulaic statement above), but also the equation of autonomy with the grounding of political critique (i.e., §2). As I argued in Chapter Four, I think we can read the analyses that Young and Foucault offer us as questioning the essentially liberal interpretation of autonomy as a concept specifiable in either procedural or substantive terms, or at some conjunction of the procedural and the substantive. Their analyses draw this framework into question because both Young and Foucault try to show, in different ways, that politics inserts itself at what we could call an earlier or more fundamental level than what is captured by the casting of autonomy in terms of the procedural/substantive distinction. If we cast autonomy in either procedural or substantive terms, or in terms of some balance between the two, we are committed to predicating over persons (or, alternatively, assigning 'persons' a place in our language-game). Young and Foucault argue that this act of predicating over persons involves a certain political investment. They argue, that is, that any 'facts' about persons are infused with value, and often infused with power.

How does this challenge the autonomy-political critique link? This link captures a basic intuition: that the more

disengaged we are from our political culture the more capable we are of criticizing the institutions that sustain this political culture. According to Kant, and the neo-Kantians, there are formal mechanisms that provide for this disengagement. For Kant, our residence in the noumenal realm provided us with the capacity to reason impartially. In terms of the grounding of political critique, for Kant, the sense in which we are members of the kingdom of ends allows us to assume a justified critical posture toward the kingdoms of the earth. In Apel's terms, for ILT_p , our membership in the unlimited communication community allows us to assume a justified critical posture toward the limited contingent communication community of which we are also members. So it is precisely the extent to which we have the capacity for autonomous judgement--a capacity guaranteed by our membership in the unlimited communication community (Apel), or our participation in dialogue approximately free from coercion (Habermas)--that marks the extent to which we can justifiably assume a critical posture to the institutions and practices that characterize the human societies in which we are members.

Young's and Foucault's arguments undermine this autonomy-political critique link by challenging the tacit assumption that the concepts of autonomy informing this link are free from the intrusion of politics. That is, Young's and Foucault's arguments suggest that politics and power are at work at the level of designating the 'persons' who participate in the dialogue at the ideal limit of freedom from coercion. The point, basically, is

this: 'autonomy' as described by neo-Kantianism cannot be linked with a disengagement from contingent political practices (and thus as a grounding of political critique), because this conception of autonomy is constrained to predicate over persons, and the process of delimiting personhood is *itself* embedded in contingent political practices.

I would argue that it is precisely because, on the one hand, the autonomy-grounding of political critique link has remained largely unchallenged, and because, on the other hand, Foucault's work challenges this link, that Foucault has been met with charges of normative confusion, accused of throwing cold water on reform, and charged with relativistic self-refutation. With his histories, Foucault tried to show us how personhood, and not simply, e.g., 'citizenship', could be historicized and politicized. No one meets the claim that 'citizenship' can be historicized and politicized with charges of relativistic self-refutation. With the exception perhaps of hard-core Lockeans, we all seem prepared to accept that our status as members of a polity is relative to time and place. I would suggest that historicizing and politicizing citizenship is non-controversial because such an analysis leaves a space for a non-relative concept of personhood. We can understand Foucault's work as an investigation into the consequences of suspending the assumption that such a gap exists between e.g., citizenship and personhood. I would argue, further, that his work succeeds in questioning the possibility of formulating a criterion of demarcation between

citizenship and personhood that is free from an investment in some economy of power. Without a non-political criterion of demarcation, the link between autonomy and grounding of political critique is fundamentally undermined. That is, if we cannot retreat to a level of personhood free from the intrusion of politics, then we cannot disengage ourselves from the institutions and practices toward which we direct our analyses.

It is this consequence of Foucault's analysis that leads many of his critics to suggest that Foucault's work is relativistically self-refuting. The response to this challenge, I would argue, is to question the tacit assumption underlying these critiques that the process of historicization and politicization has boundaries beyond which it cannot legitimately pass. That is, why is it that we can accept the historicization and politicization of 'citizenship', but draw a line at the point of 'personhood'? Why does the logical claim of the self-refuting status of relativism apply at the latter point and not the former? I suspect the self-refutation charge is motivated exactly by the assumption that Foucault is drawing into question: that there is a level of personhood prior to and independent of citizenship. Now, at a certain level, this response is irrefutable. Foucault's analysis is entirely compatible with any purely ontological theory of personhood. Here, oddly enough, I think Rorty comes to Foucault's rescue. Recall, as I covered in Chapter Three, that Rorty argued for the adoption of a methodological anti-essentialism with respect to personhood in

the domain of political theory. In short, Rorty argued that, for the purposes of normative political theory, 'apparent persons' are as good as 'real persons.' The stronger point is that, even if there are such things as 'real persons', no politically normative consequences follows from the nature of such 'real persons.' That is, it is my intuition--which I cannot defend here--that any ontological theory of personhood is compatible with any normative political theory. The burden falls on Foucault's critics who must argue for two claims: (1) we can establish, once and for all, the ontological status of personhood, and (2) we can demonstrate how some particular normative political theory follows with deductive security from the ontological status of personhood. My intuition--which, again, I cannot defend here--is that neither of these claims can be conclusively settled.

In sum, I am arguing that the critics who charge Foucault with relativistic self-refutation base their criticism on (1) placing an arbitrary boundary beyond which the program of historicizing and politicizing cannot legitimately pass (i.e., accepting the relativizing of citizenship but not personhood), and/or (2) holding out for the 'apparent persons/real persons' distinction by arguing that the ontological status of personhood can both be determined, and have direct consequences for normative political theory. I am defending Foucault by arguing that (1) there is no way to justify halting the process of historicization and politicization before the level of

personhood, and (2) even if the ontological status of personhood could be determined--which I doubt--, it would have no implications for normative political theory.

I am not suggesting, however, that Foucault's analysis does not have serious consequences for normative political theory. Indeed, it follows from my reading of Foucault that the neo-Kantian grounding of political critique that I defended in this thesis is strongly undermined by Foucault's (and Young's) challenges to the neo-Kantian account of autonomy. But I do not think that the neo-Kantian tradition must thus be abandoned. I would suggest that the most important problem that emerges from the current dialogue in contemporary political theory involves the reconciliation of the insights of the neo-Kantian tradition with the challenges offered by Foucault and Young. That is, I think that the procedural interpretation of ideal limit theory *does* provide us with a non-arbitrary starting point for normative political theory. However, Foucault and Young succeed in demonstrating that this neo-Kantianism has at its core an account of autonomy that is in need of serious revision. The task is, therefore, to neither abandon the program of articulating a defensible grounding for political critique, nor to reject the challenges that show us how much the attempts to do so have so far been unable to disengage themselves from contingent political

practices and structures of power.¹ More specifically, political theorists must come to terms with the consequences that a thorough-going politicization of personhood has for the concept of autonomy and the critique of ideology.

¹Stephen K. White has already outlined what such a research program might look like. See his "Poststructuralism and Political Reflection", Political Theory 16:2 (1988), pp. 186-207.

APPENDIX: KANT DEFENDED AGAINST STROUD AND RORTY

As I indicated above, an earlier draft of this thesis included a defence of Kant against Stroud's 'verification principle' criticism and Rorty's argument that TAs require a scheme-content distinction and faith in privileged access to one's own subjectivity. These passages, however, are rather parenthetical with respect to my overall argument. So in the interests of preserving a clear narrative current to my paper, I have moved them into this appendix.

§ 1 Recall that Stroud argued that transcendental arguments cannot satisfy what Genova called the 'objectivity thesis'--in Kant's words, the demonstration that the "*subjective conditions of thought can have objective validity*" (A 89/B 122)--without recourse to a covert verification principle. That is, Stroud argued, we can only get from our use of concepts (meaning) and the conditions whereby those concepts obtain in the world (truth) if we make some essential link between meaningfulness and truth in a way that securing the conditions of the former means securing the conditions of the latter. This link, Stroud argues, is a form of the verification principle.

In Genova's terms, Stroud is arguing that the move from a metaphysical deduction (MD) to a transcendental deduction (TD) requires a covert verification principle. An MD "provides a priori justification of a unique [conceptual framework]--a

[conceptual framework] which is a necessary presupposition of all possible contingent interpretations of experience." (Genova, pp. 479-480). A TD establishes the objectivity thesis, arguing that this unique conceptual framework is objectively valid. Stroud argues that the sceptic is free to intervene at the point of transition from the MD to the TD. He argues the MD-TD move can be accomplished only by smuggling in a verification principle.

As Genova points out, Kant was obviously aware of the gap that Stroud points to (and that Stroud argues is filled with a tacit verification principle). The MD/TD distinction that Genova draws come from Kant.¹ Yet Kant often speaks as though there were no gap to fill. In the first edition version of the *Transcendental Deduction*, Kant claims,

[the concepts] which contain a *a priori* the pure thought involved in every experience, we find in the categories. If we can prove that by their means alone an object can be thought, this will be a sufficient deduction of them, and will justify their objective validity. (A 96-A 97)

In terms of the distinctions introduced above, Kant seems to be moving from an MD (the claim that these concepts are *a priori* and unique), based on an HD (the conditional claim that these

¹Genova points to two other distinction in Kant that parallel the MD/TD distinction: Kant's two characterizations of an object (any intentional object of consciousness as opposed to the epistemic object that stands under the schematized categories), and Kant's contrast between an objectively underdetermined experience and an objectively determined experience (Genova, p. 479). These distinctions draw on material in the first Critique--primarily the section on "Schematism"--with which I have not dealt, so I will not discuss this further substantiation of the claim that Kant was well aware of the gap to which Stroud's critique points.

concepts are necessary for thought), to a TD (the claim that these objects enjoy "objective validity"). Kant appears to argue that providing vindication of these concepts by an MD ("If we can prove that by their means alone an object can be thought") will simultaneously achieve the goals of a TD ("...and will justify their objective validity"). This argument thus seems to be committing the same sort of equivocation that Stroud argued must be grounded in a covert verification principle.

Genova argues that, in Kant, Stroud's gap is filled not by covert verification principle, but rather by the 'metaphilosophical context' of Kant's Copernican Revolution. "The reason why Kant can make this logical connection [between the subjective necessity of the categories and their objective validity]," Genova argues, "is that his Copernican principle provides Kant a basis on which he can establish a new criterion of objectivity" (Genova, p. 487). The hypothesis that Kant transfers via analogy from (what he takes to be) the methodology of mathematics and physics--that "we may have more success in the tasks of metaphysics, if we suppose that objects must conform to our knowledge", rather than supposing that our knowledge must conform to objects--implies that the notions of objects and objectivity must shift away from metaphysical realism.

What is to be construed as an object of knowledge is not something that is given independently of our cognitive activity. If that were the case, a priori knowledge [which, given the Copernican principle, is to be the object of the inquiry into the conditions of experience] would be out of the question because no necessary truths could be derived from the experience of such independent objects...When representations

represent an object (as opposed to being merely subjective), there is not thereby a determinable conformity between the representation and some non-epistemic transcendent thing, but a determinable nexus between the representations and other representations in conformity with universal rules.² (Genova, p. 487)

Within this perspective, Kant's demonstration that certain concepts are necessary for experience and have a priori validity (that is, are not derived from empirical generalization--as Hume thought 'causality' was) will suffice to close the gap between subjective necessity and objective validity, because--according to the Copernican principle--there can be no further intelligible appeal to an appearance/reality hypothesis. This further appeal would have to evoke things-in-themselves as distinct from things as they appear to us, and, given that the boundaries of knowledge are co-extensive with the boundaries of experience, we can have no knowledge of things-in-themselves. Thus, such an appeal is non-sensical.

How satisfying is Kant's response (or, rather, Genova's response to Stroud on Kant's behalf)? An important point to make is that, as Genova notes, Kant's Copernican principle is not equivalent to a VP.

The [Copernican] principle, as we have seen, postulates that an object of knowledge is not something that can be given independently of our cognitive activity--but

²This point serves to foreshadow a perspective I will defend later in this chapter: that procedural ILT is a direct descendant of the critical philosophy. ILT philosophers ask us to abandon truth as correspondence for truth as rational assertability, and thus make objectivity a function of the *rules* of consensus-formation (in a manner analogous to Kant's connection of objectivity with the rules of thought given by the a priori concepts of understanding).

this is not equivalent to saying that one could not have a significant belief without knowing how, in principle, the belief could be verified or falsified; or that one could not have a significant concept without knowing how, in principle, to determine whether or not the concept could be instantiated. (Genova, p. 492)

Genova argues that the Copernican principle does not appear as a covert premise in Kant's TD, but rather as a background meta-philosophical context. The legitimacy of Kant's TD thus falls to the question of whether this background principle can be shown to be a philosophical liability (Genova, p. 495).

It not at all clear how we are to adjudicate over such Big-Picture-level questions as whether the Copernican principle is a philosophical liability. Kant, of course, had a criterion in mind. Recall that he supported his endorsement of the Copernican principle by appeal to the success that he argued mathematics and physics enjoyed by their employment of this principle. So something like 'greater explanatory success' appears to have been his overriding concern. Such a goal or virtue, however, is itself theory-laden: we need criteria to unpack the terms 'greater' and 'success' in this overriding criterion. At this level of analysis, as well, it is appropriate to question Kant's philosophy of science. We may well be prepared to agree with his assessment of mathematics,³ but contemporary philosophy of science hardly takes physics to be an a priori discipline. As

³That is, his assessment of mathematics as a priori, not as exclusively Euclidean. I don't think the move to non-Euclidean geometry marked a move in mathematics to the a posteriori.

such, it may be argued that Kant's argument for the Copernican hypothesis breaks down at the point of his assessment of the methodology of natural science. Indeed, one of the most obvious problems with the first Critique is that the Transcendental Deduction accords a priori status to a fundamentally Newtonian picture of the world. Does it follow from our abandonment of Newtonian physics that we ought also to abandon the transcendental a priori?

There seem to be two responses to this problem. The first, advocated primarily by Jay Rosenberg,⁴ is that we replace Kant's idea of an ahistorical conceptual core with core pluralism. However, this revision only begs the transcendental question, as it leaves unanswered the issue of how it is we can generate conceptual schemes in the first place, i.e., it does not address the question of the precondition of experience (or perhaps, the preconditions of constructing theories about experience). The other alternative, which I have defended, is a revision of Kant that recasts the transcendental questions in procedural terms: what I have called the procedural interpretation of ideal limit theory.

The ideal limit theory revision of Kant, however, leaves unanswered the big picture question that was raised above. ILT_p asks that we abandon correspondence theories of truth for a 'rational assertability' theory of truth (i.e., a proposition is

⁴Cf. Jay Rosenberg, "Transcendental Arguments Revisited", Journal of Philosophy 72 (1975), pp. 611-624.

true iff it is rational to assert it). Like Kant, ideal limit theorists propose epistemological theories of truth. While, as I argued above, Kant did offer some justification for his epistemological theory of truth--i.e., the science-analogy appeal for the Copernican hypothesis--it strikes me that his approach does not stand or fall on the defensibility of this justification. That Kant was wrong about science--as we would claim today--does not mean that he was also incorrect about the basic structures of knowledge, but only that this appeal no longer has the force it had in Kant's day. The strongest argument in favour of an epistemic theory of truth (i.e., that 'truth' is a tag we put on justified epistemic practices) would be a refutation of a metaphysical theory of truth (i.e., that 'truth' is a function of how our epistemic activities match up with the world). Now, while Kant offered some arguments against the latter conception of truth (cf. B xx), and Putnam defends his version of ideal limit theory after purportedly refuting metaphysical theories of truth⁵, I do not think this issue can be conclusively settled. Any epistemic theory of truth will be logically compatible with a metaphysical theory of truth. Even if it is argued--as Kant and Putnam argue--that we cannot make sense of truth independent of our epistemic activities, this claim is still compatible with the existence of a reality independent of our thinking and practices with which our picture

⁵Cf. Putnam, Reason, Truth, and History (Cambridge: Cambridge University Press, 1981).

of the world agrees to a greater or lesser degree. Successful arguments of Kant's and Putnam's type can *persuade* us to abandon metaphysical realist theories of truth because such theories lack any utility, but they cannot conclusively defeat the realist position. This does not mean that the Kant-ideal limit theory account is *incorrect*, but that it cannot have final say in the metaphysical realist's terms.

Does this mean that Stroud's gap is reopened in the transcendental project? Perhaps, but at the very abstract metaphilosophical level of the realism issue. As Genova argues, the sceptical gap in Kant can be opened only with the rejection of the Copernican principle. The defensibility of this rejection seems to turn on a question of burden of proof. From the realist/sceptic's position, the burden falls on Kant to show why we ought to adopt the Copernican principle and abandon the metaphysical realist alternative. From Kant's perspective, the burden falls on the sceptic/realist to defend why we should maintain the realist position (or, in the case of the sceptic, not be content until the realist position is defeated on the realist's turf) when this position does not capture our epistemic practices (as Kant and Putnam argue). I do not think that this issue admits of resolution. However, I would argue that the adoption of an epistemological theory of truth--for Kant, the Copernican principle, for ILT_p, a 'rational assertability' theory of truth--involves issues large enough and of such apparent lack of resolvability that the *assumption* of such a theory of truth is

not irresponsible.

So we can say, finally, that Stroud was correct in arguing that the sceptic is not fully answered by the claims of a TA, but that the gap to which the sceptic's doubts may take recourse occurs further up the levels of philosophical abstraction than was suggested by Stroud's claim that TAs rest on covert verification principles.⁶

§ 2 Rorty argues that Kantian TAs presuppose a distinction between scheme and content, "e.g., one between concepts and intuitions, or thought and the objects of thought, or words and the world" (TASRP, p. 79). This presupposition is shared with other forms of argumentation; for an argument to be a Kantian TA, it must also satisfy the following conditions:

[a] The scheme-content distinction is construed as a distinction between that which is better known to us (our subjectivity, roughly) and that which is less known to us.

[b] Our 'legitimizing' transcendental knowledge of the necessary truth that content will correspond to scheme, is made possible by the fact that our subjectivity (the scheme) creates the content.⁷
(TASRP, p. 79)

Rorty argues that Kant's TA thus becomes problematic when one becomes dubious about the notion of privileged access to one's own subjectivity (TASRP, p. 80).

⁶'Higher up', at least, for Kant and ideal limit theorists. Stroud's criticism may well be sound with respect to Strawson and Shoemaker, but this question is outside of my current focus.

⁷according to Kant's Copernican principle.

I would argue that Rorty's analysis of Kant is incorrect. While the heart of Kant's epistemology was an explanation of the necessity of organizing experience under categories, it does not seem to me that his task was analogous to trying to articulate the correspondence between, e.g., "thought and the objects of thought, or words and the world." While in many passages of the first Critique⁸ Kant argues that experience requires something other than the agent experiencing--"otherwise we should be landed in the absurd conclusion that there can be appearance without anything that appears" (B xxvi)--, it was not Kant's goal to explain how experience linked up with the world. On the contrary, Kant argued that, while we are forced to *think* of objects-in-themselves--else we fall into contradiction like the "absurd conclusion" mentioned above--we cannot *know* these objects. Further, Kant argued, the supposition that we can know these objects, that we can try to fit our scheme to the world, lands us in Humean scepticism. As Genova argues in a passage cited above, by the Copernican principle 'objectivity' takes on a meaning quite different from what we find in realism: "When representations represent an object (as opposed to being merely subjective), there is not thereby a determinable conformity between the representation and some non-epistemic thing, but a determinable nexus between the representations and other representations in conformity with universal rules." (Genova, p.

⁸primarily the Preface to the Second Edition, the Analogies, the Transcendental Deduction and the Refutation of Idealism.

487). Now Rorty does not get Kant wrong here, but I think he is forcing Kant into an inappropriate conceptual slot. Rorty does not impute a traditional correspondence theory of truth to Kant, but argues that, according to Kant, "[o]ur 'legitimizing' transcendental knowledge of the necessary truth that content will correspond to scheme, is made possible by the fact that our subjectivity (the scheme) *creates* the content." Now while this account is certainly accurate given a particular interpretation of scheme and content, I am suggesting that lumping Kant's epistemology together with traditional realist epistemologies--wherein scheme/content is interpreted as "thought/the objects of thought" or "word/world"--obscures important differences between Kant's transcendental idealism and this type of realism. Most fundamentally, this latter form of realism seems to imply a metaphysical theory of truth, while Kant's epistemology--I have argued--is a purely epistemological theory of truth.

Nor do I think that Rorty's claim that Kantian TAs depend upon the notion of privileged access to one's own subjectivity is quite accurate. To be sure, Kant's Copernican principle asks us to turn inward to discover the conditions of experience and the criteria of legitimate epistemic judgement. But Kant is not like Descartes (a tradition in which, in this context, Rorty places him): Kant does not think we have certainty about the objects of introspection. Kant does not try to build knowledge of the world from the austere self-certifying propositions at which we can arrive through self-reflection. Kant's recourse to the activity

of reason in the production of knowledge was not occasioned by a faith, on Kant's part, as to the infallibility of, say, what the positivists called 'brute sense datum reports.' Kant's Copernican principle asks us to understand objects as conforming to thought because an epistemology that assumes that thought conforms to objects ends up in contradictions and irresolvable scepticism. Furthermore, history indicated to Kant that while metaphysicians could find no agreement on any issue, mathematics and physics were on the road to progress. This progress, Kant argued, was made possible by the assumption of what he introduced into metaphysics as his Copernican principle. Kant's Transcendental Deduction is not based on a faith in the infallibility of self-reflection, but on the claim that we can provide a legitimation of knowledge only if we begin with the activity of the knower.

So by arguing that Kant's TA requires a scheme-content distinction and faith in privileged access to own's own subjectivity, Rorty is obscuring important differences between Kant and metaphysical realists on the one hand, and between Kant and Cartesians on the other.

Bibliography

- Allen, Robert Van Roden, "Emancipation and Subjectivity: A Projected Kant-Habermas Confrontation", Philosophy and Social Criticism 9:3-4, pp. 283-303.
- Allison, Henry E., "Practical and Transcendental Freedom in the Critique of Pure Reason", Kant-studien 73 (1982), pp. 271-290.
- Apel, Karl-Otto, "The Problem of Philosophical Foundations in Light of a Transcendental Pragmatics of Language", Kenneth Baynes, James Bowman, and Thomas McCarthy (eds.), After Philosophy (Cambridge: MIT Press, 1987), pp. 250-290.
- _____, Towards a Transformation of Philosophy (London: Routledge & Kegan Paul, 1980).
- Beatty, Joseph, "The Rationality of the 'Original Position': A Defence", Ethics 93 (1983), pp.484-495.
- Beck, Lewis White, A Commentary on Kant's Critique of Practical Reason (Chicago: University of Chicago Press, 1960).
- Cornell, Drucilla, and Adam Thurschwell, "Feminism, Negativity, Intersubjectivity", Praxis International 5:4 (1986), pp. 484-504.
- Darwall, Stephen L., "A Defense of the Kantian Interpretation", Ethics 86: 2 (1976), pp. 164-170.
- Delue, Steven M., "Kant's Politics as an Expression of the Need For His Aesthetics", Political Theory 13:3 (1985), pp. 409-429.
- Douglass, Bruce, "The Common Good and the Public Interest", Political Theory 8:1 (1980), pp. 103-117.
- Ewing, A.C., A Short Commentary on Kant's Critique of Pure Reason (London: Methuen & Co., 1938).
- Flathman, Richard E., "Review of Contingency, Irony, and Solidarity", Political Theory 18:2, pp. 308-312.
- Fraser, Nancy, "Foucault on Modern Power: Empirical Insights and Normative Confusions", Praxis International 1 (1981), pp. 272-287.
- Foucault, Michel, Discipline and Punish (New York: Vintage Books, 1979).

- _____, "The Ethic of Care for the Self as a Practice of Freedom", James Bernauer and David Rasmussen (eds.), The Final Foucault (Cambridge: MIT Press, 1988), pp. 1-20.
- _____, Foucault Live, Sylvère Lotinger (ed.) (New York: Semiotext(e), 1989).
- _____, The History of Sexuality (Vol. 1) (New York: Vintage Books, 1980).
- _____, Power/Knowledge, Colin Godon (ed.), (New York: Pantheon Books, 1980).
- _____, "The Subject and Power", in Hubert Dreyfus and Paul Rabinow, Michel Foucault: Beyond Structuralism and Hermeneutics (Chicago: Chicago University Press, 1982), pp. 208-226.
- _____, "Questions of Method", Baynes, et.al. (eds.), pp. 100-117.
- _____, "What Is Enlightenment?", Paul Rabinow (ed.), The Foucault Reader (New York: Pantheon, 1984), pp. 32-50.
- Genova, A.C., "Good Transcendental Arguments", Kantstudien 75, pp. 469-495.
- Gram, Moltke S., "Must Transcendental Arguments Be Spurious?", Kantstudien 65 (1974), pp. 304-317.
- _____, "Must We Revisit Transcendental Arguments?", Philosophical Studies 31 (1977), pp. 235-248.
- _____, "Transcendental Arguments", Nous 5 (1971), pp. 15-26.
- Grimshaw, Jean, Philosophy and Feminist Thinking (Minneapolis: University of Minnesota Press, 1986).
- Habermas, Jürgen, Communication and the Evolution of Society (Boston: Beacon Press, 1979).
- _____, "The Hermeneutic Claim to Universality", Josef Bleicher (ed.), Contemporary Hermeneutics (London: Routledge & Kegan Paul, 1980), pp. 181-211.
- _____, Knowledge and Human Interests (Boston: Beacon Press, 1971).
- _____, Theory and Practice (Boston: Beacon Press, 1973).
- _____, The Theory of Communicative Action (Boston: Beacon Press, 1981).

- _____, Toward a Rational Society (Boston: Beacon Press, 1970).
- Hacking, Ian, "The Archeology of Foucault", David Couzens Hoy (ed.), Foucault: A Critical Reader (London: Basil Blackwell, 1986), pp. ?-?.
- Hiley, David R., "Foucault and the Analysis of Power: Political Engagement Without Liberal Hope of Comfort", Praxis International 4:2 (1984), pp. 192-207.
- Hintikka, Jaakko, "Transcendental Arguments: Genuine and Spurious", Nous (1972), pp. 274-281.
- Johnson, Oliver A., "Autonomy in Kant and Rawls: A Reply", Ethics 87 (1976), pp. 251-254.
- _____, "The Kantian Interpretation", Ethics 85 (1974), pp. 58-66.
- Kant, Immanuel, The Critique of Judgement, trans. J.C. Meredith (Oxford: Oxford University Press, 1952).
- _____, Critique of Practical Reason, trans. Lewis White Beck (New York: Macmillan Publishing Co., 1956).
- _____, Critique of Pure Reason, trans. Norman Kemp Smith (London: Macmillan Education Ltd., 1929).
- _____, Groundwork of the Metaphysics of Morals, trans. H.J. Paton (New York: Harper & Row, 1964).
- _____, Prolegomena to Any Future Metaphysics, trans. L.W. Beck (Indianapolis: Bobbs-Merrill, 1950).
- _____, Religion Within the Limits of Reason Alone, trans. T.M. Greene and H.H. Hudson (New York: Harper & Row, 1960).
- Kelly, George Armstrong, Idealism, Politics and History (Cambridge: Cambridge University Press, 1969).
- McCarthy, Thomas, The Critical Theory of Jürgen Habermas (Cambridge: MIT Press, 1978).
- Neal, Patrick, "Justics as Fairness: Political or Metaphysical?", Political Theory 18:1 (1990), pp. 24-50.
- Nedelsky, Jennifer, "Reconceiving Autonomy", Yale Journal of Law and Feminism 1:1 (1989), pp. 7-36.
- Norris, Christopher, The Contest of Faculties (London: Methuen & Co. Ltd., 1985).

- Oliver, Leaman, "Transcendental Arguments: Gram's Objections", Kantstudien 69 (1977), pp. 468-477.
- O'Neill, Onora, "The Public Use of Reason", Political Theory 14:4 (1986), pp. 523-551.
- Peirce, Charles Saunders, Philosophical Writings of Peirce, ed. Justus Buchler (New York: Dover, 1955).
- Putnam, Hilary, The Many Faces of Realism (Illinois: Open Court, 1987).
- _____, Reason, Truth, and History (Cambridge: Cambridge University Press, 1981).
- Rawls, John, A Theory of Justice (Cambridge: Belknap Press, 1971).
- _____, "Kantian Constructivism in Moral Theory", Journal of Philosophy 77:9 (1980), pp. 515-5??.
- Reiss, Hans (ed.), Kant's Political Writings (Cambridge: Cambridge University Press, 1970).
- Riley, Patrick, Kant's Political Philosophy (New Jersey: Rowman and Littlefield, 1983).
- Rorty, Richard, Consequences of Pragmatism (Minneapolis: University of Minnesota Press, 1982).
- _____, Contingency, Irony, and Solidarity (Cambridge: Cambridge University Press, 1989).
- _____, "Postmodernist Bourgeois Liberalism", Robert Hollinger (ed.), Hermeneutics and Praxis (Notre Dame: University of Notre Dame Press, 1985), pp. 214-221.
- _____, "Solidarity or Objectivity?" John Rajchman and Cornel West (eds.), Post-Analytic Philosophy (New York: Columbia University Press, 1985), pp. 3-19.
- _____, "Transcendental Arguments, Self-Reference, and Pragmatism", Peter Bieri, Rolf P. Horstmann, and Lorenz Kruger (eds.), Transcendental Arguments and Science (Holland: D. Reidel, 1979), pp. 77-104.
- _____, "Verificationism and Transcendental Arguments", Nous 5 (1971), pp. 3-14.
- Rosenberg, Jay F., "Transcendental Arguments and Pragmatic Epistemology", Bieri, et.al. (eds.), pp. 245-261.

- _____, "Transcendental Arguments Revisited", Journal of Philosophy 72 (1975), pp. 611-624.
- Saner, Hans, Kant's Political Thought (Chicago and London: Chicago University Press, 1973).
- Stroud, Barry, "Transcendental Arguments", The Journal of Philosophy 65:9 (1968), pp. 241-256.
- _____, "Transcendental Arguments and 'Epistemological Naturalism'", Philosophical Studies 31 (1977), pp. 105-115.
- _____, "The Significance of Scepticism", Bieri, et.al. (eds.), pp. 277-297.
- White, Stephen K., "Poststructuralism and Political Theory", Political Theory 16:2 (1988), pp. 168-208.
- _____, The Recent Work of Jürgen Habermas (Cambridge: Cambridge University Press, 1988).
- Williams, Howard, Kant's Political Philosophy (Oxford: Basil Blackwell, 1983).
- Young, Iris Marion, "The Ideal of Community and the Politics of Difference", Social Theory and Practice 12:1 (1986), pp. 1-26.
- _____, "Impartiality and the Civic Public: Some Implications of Feminist Critiques of Moral and Political Theory", Praxis International 5:4 (1986), pp. 381-401.