

Consensus Between Discourse Ethics and Democracy: Habermas on Compromise

Nicholas Zavediuk

A Thesis
in
The Department
of
Philosophy

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Arts (Philosophy) at
Concordia University
Montréal, Québec, Canada

August 2005

© Nicholas Zavediuk, 2005



Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 0-494-10178-4

Our file Notre référence

ISBN: 0-494-10178-4

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

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

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

ABSTRACT

Consensus Between Discourse Ethics and Democracy: Habermas on Compromise

Nicholas Zavediuk

The following thesis is a critical consideration of the role compromise formation plays in both Jurgen Habermas's discourse ethics and his discourse theory of democracy. In particular, I want to demonstrate how deficiencies his critics have identified in his political theory—notably the strong bifurcation between the informal public sphere and the formal parliamentary-administrative complex—can be traced back to problems associated with his moral theory, particularly its “constitutive presupposition” that engaging in practical discourse will bring about an argumentatively achieved consensus on a valid norm.

My argument is that compromise need not be, as Habermas thinks, a strategic form of action coordination. Drawing on Joseph Heath's rational choice model of nonstrategic bargaining and a ‘defeasibility’ reading of the universalization principle in discourse ethics, I hope to establish that ‘reasonable compromises’ are capable of discursive moral justification.

The political consequences of this investigation will focus on how an enhanced understanding of moral compromise formation can expand the role of the public sphere in deliberative democratic theory (something of a reply to Habermas's critics) while retaining the normative insights of discourse ethics.

Acknowledgements

I would first like to thank my thesis supervisor, Dr. Matthias Fritsch, whose constructive comments, enthusiasm for my project, and deep knowledge of the current research surrounding Habermas's work helped me immensely. A special debt is owed to James Gifford and Lindsay Parker for their editing talents and sympathetic ears. My friendships in Montréal (Cecilie, Petey, Paulo, Chris, Alex, Terry, the Java U-Guy crew, and many of the other graduate students and professors at Concordia University too numerous to mention) enriched not only my academic experience but also provided an inspirational milieu for a remarkable 18 months of my life. Finally, none of this would have been possible without the continued support of my family: Orest, Tom and Ainsley.

I dedicate this thesis to the memory of my mother, Carla, and all of the wonderful lessons she taught me and traits she instilled in me.

Table of Contents

Introduction:	1
Chapter One: Law and Democracy in <i>Between Facts and Norms</i>	8
1.1	8
1.2	14
1.3	19
Chapter Two: Deliberation, Discourse and Consensus	31
2.1	33
2.2	43
2.3	54
Chapter Three: Discourse, Compromise and the Public Sphere	66
3.1	68
3.2	76
3.3	84
Conclusion	94
End Notes	96
Works Cited	98

Introduction

“The fundamental idea of democratic, political legitimacy,” Joshua Cohen writes, “is that the authorizations to exercise state power must arise from the *collective decisions* of the equal members of society who are governed by that power” (185). By and large, theories of deliberative democracy seek to understand how reasoned discussions between free and equal citizens in civil society can come to influence or shape political decision-making. For deliberative democrats, in one way or another and to greater and lesser degrees, this political legitimacy is not so much tied to power or stability, and collective decisions are not so much tied to authorization by vote and plebiscite, as they are both tied to the reason and will of the citizenry. However, characterizations like this say very little about how deliberative ideals are to work in practice, and even less about how deliberation actually connects with expressions of political power. For example, in most representative democracies, elected officials usually profess that they *do* listen to what the people say, and that the opinions of their constituents *do* matter when they cast votes in whatever formal political body they have been elected to. And town hall meetings, where these constituents invite their representative to listen to the opinions and arguments of the electorate on this or that issue, are intended to tie the authorized decision-maker with those that are eventually bound by whatever decisions they make. The deliberative democrat wants to go beyond these vague associations, which are as likely to fuel political apathy or cynicism in an electorate as empowering the electorate:

According to the deliberative interpretation of democracy, then, democracy is a system of social and political arrangements that institutionally ties the exercise of power to free reasoning among equals. (Cohen 193)

Of course, what are as important as the ideals are the details, and in giving these details, deliberative democrats have to find a way to reconcile their normative impulses with the empirical exigencies of modern societies that threaten to occlude their realization. These exigencies include value pluralism within the citizenry, as well as the increasing complexity of societies with market economies, welfare states, and extensive legal structures for mediating and constraining social relationships. Two questions become immediately obvious from the above description of the deliberative interpretation of democracy. One, how can deliberations in complex societies be 'institutionally tied' to political decisions, and two, what exactly is meant by 'free reasoning among equals?'

One study that has been particularly influential in its approach to describing the relationship between formal political bodies and the informed debate of the citizenry is Jürgen Habermas's discourse theory of democracy. In the discourse-based model of democracy, Habermas incorporates insights from his theory of communicative rationality to permeate the seeming incongruity between the normative meaning of law and modern political institutions. Generally speaking, legitimate laws are those that command, or could potentially command, the "justified assent" of all subject to them. Yet, in complex democratic societies, the task of creating laws and policy increasingly falls to representative parliamentary bodies, institutions that are both at considerable remove from the citizens they ostensibly represent and whose decisions are invariably the product of majority voting practices. For Habermas, legitimate law and democratic rule:

Are compatible only if the latter has an internal relation to the search for truth: public discourse must mediate between reason and will, between the opinion-formation of all and the majoritarian will-formation of the representatives. (Habermas, "Popular Sovereignty" 475)

In the discourse theory of democracy, the pragmatic presuppositions of a practical discourse, defined simply as an intersubjective communicative practice based on “the *idealizing* assumptions that everyone who seriously engages in argumentation must make *as a matter of fact*,” assume a modified public form as “the institutionalization of the corresponding procedures and conditions of communication, as well as the interplay of institutionalized deliberative processes with informally constituted public opinions” (Habermas, Justification 50; Facts and Norms 298).

While Habermas makes it clear that positive law and his own discursive moral theory of communication have only a “complementary” relation to each other, there is a strong correspondence between certain ideals in Habermas’s communicative ethics and his ‘democratic principle,’ which states: “that only those statutes may claim legitimacy that can meet with the assent of all citizens in a discursive process of legislation that in turn has been legally constituted” (Facts and Norms 110). However, discourse ethics is a highly demanding theory of moral argumentation, and developing a democratic theory based partly on the idealizations of undistorted communication turns on whether or not these idealizations—especially the equation of rationally motivated agreement with an argumentatively achieved *consensus* on validity—can be convincingly used as the normative underpinning for actual political institutions under conditions of “contemporary value pluralism.” Even in recognizing that “practical discourses are not a blueprint for institutions, but they can help evaluate prevailing institutional arrangements” (Benhabib, Claims 107), my opinion is that any engagement with Habermas’s political theory cannot underplay the effect his moral theory exercises over his democratic proceduralism.

The thesis I want to advance here is that a broader understanding of compromise formation at the level of moral theory may be used to mitigate the force of certain criticisms that have been directed at Habermas's political theory. For a variety of reasons, these criticisms reproach Habermas for the way his 'two-track' model of democracy, which divides political power between a robust civil society and the formal administrative-parliamentary complex, constrains and diminishes the influence of the public sphere on actual political decision-making. My argument is that if reasonable compromises are capable of being justified in a practical discourse (contrary to Habermas's own view), then this weakens the basis for the strong separation Habermas makes between the public sphere and authoritative political institutions and, in so doing, creates a space to empower the public sphere without sacrificing the normative insights of discourse theory.

My approach is motivated by four concerns that will be incorporated into my thesis. First, certain institutional features of Habermas's deliberative political model, especially the demarcation of the public sphere from state power, can be partially read as an effort to lessen what Habermas sees as the corrosive effects of bargaining and compromise on the consensual creation of communicative power. The problem is that fashioning "too strong a distinction between will-forming and opinion-forming institutions undermines any actual democratic sovereignty" (Bohman, Public Deliberation 185). Second, interpreting the function Habermas assigns to compromise formation in his democratic theory requires foremost a familiarity and engagement with his theory of moral justification and its definition of 'rationally motivated agreement.' As Thomas McCarthy points out, Habermas views "bargaining and compromise as

reasonable means for dealing with conflict,” but “he subordinates them to the achievement of consensus in practical discourse” (McCarthy 189). Third, even though Habermas believes “compromises make up the bulk of political processes,”¹ he thinks they cannot be “conducted in the form of a rational discourse that neutralizes power and excludes strategic action” (“Three Normative” 245). Moreover, by restricting the discourse-theoretic interpretation of compromise formation to just the clarification of “fair bargaining conditions” and the “balance of interests,” Habermas overlooks the potential for compromise to “integrate the various interests and ideals in play, and to reach solutions that are mutually acceptable and embody equal concern and respect for those involved” (Bellamy 111). Fourth, while there is a wide difference of opinion amongst democratic theorists regarding the impact value pluralism has on the stability and inclusiveness of democratic institutions, it seems to me that any degree of value pluralism necessarily modifies the *kind* of agreement that democratic arrangements should seek to bring about. Consider Amy Gutmann and Dennis Thompson’s argument that persistent disagreement in democratic societies reflects neither an unwillingness on the part of citizens to be self-critical of their own values nor a lack of interest in seriously considering the point of view of others but instead exposes a gap between a common interest in reaching agreement and a unified understanding of that interest that cannot be brought together in any politically effective way. I believe this perspective does not invalidate the prospect that communicative ethics can provide a normative foundation on which to assess existing democratic procedures and institutions, rather:

It simply expresses a recognition that any particular moment in history we cannot collectively resolve some moral dilemmas on their own terms. Some dilemmas may have a uniquely correct solution, and some may not. But at any historical moment our imperfect understanding manifested in

the fundamental disagreements among the most thoughtful and good-willed citizens, prevents us from distinguishing those that do from those that do not. We live in a state of moral conflict over many issues about which we do not (yet) know the right answer. (Gutmann and Thompson 25)

This statement captures the central challenge that broad value pluralism imposes on deliberative democratic ideals: that unanimous agreement on political matters is unlikely under these conditions and that moral conflict should be treated as a permanent feature of political life. The permanency of deep moral disagreement signifies a need to explore approaches to democratic agreement that do not necessarily view unanimous consent on the outcome as a priority.

This thesis consists of three chapters. Chapter one is mainly expository and tries to bring out the more significant themes in Habermas's principal political treatise, Between Facts and Norms. Although this chapter tries to touch on many of the diverse ideas Habermas presents in the book, I am primarily concerned with the relationship he establishes between the political public sphere and formal decision-making bodies. Chapter two introduces several criticisms specifically directed at Habermas's political theory. While some of these criticisms can be deflected by scrutinizing Habermas's beliefs regarding the functional differentiation of society into systems and lifeworld, all of them seem to warrant closer consideration of how the discourse model of democracy assimilates certain prominent features of discourse ethics—above all the “constitutive presupposition” (McCarthy) of rational consensus—into its theoretical framework. Based on a critique of the universalization principle in discourse ethics developed in chapter two, chapter three confronts Habermas's narrow approach to compromise as a mechanism for democratic conflict resolution. More specifically, this chapter will incorporate arguments from Joseph Heath and Arash Abizadeh to illustrate how

reasonable compromises are compatible with the practice of moral justification in practical discourses. The thesis will conclude by looking at how some of these different observations on compromise formation have found a workable political expression (an expression that acknowledges Habermas's moral and political framework) in James Bohman's model of moral compromise.

Chapter One: Law and Democracy in *Between Facts and Norms*

1.1.

In *Between Facts and Norms*, Jurgen Habermas explores the association between law and democracy in large, complex and pluralistic societies. Although the book involves many different strands of historical, legal, sociological and moral thought, it can best be understood as a model of political legitimacy that employs the normative resources of Habermas's own discourse theory to more clearly understand the relationship between civil society, legal structures, and political institutions. In drawing out and assessing more precisely the nature of this connection, Habermas fashions a normative standard of political praxis that, by asking whether "the interdependent relationship between law and democracy required if either is to be legitimate" (Baynes, "Introduction" 2), can be drawn out more fully in order to critically evaluate already existing democratic and legal arrangements.

Habermas's analysis starts with a fundamental assertion: the medium of law has a dual character. The first aspect of this dual character is the *facticity* of law, which refers to the idea of positive law as both a constitutionally-protected set of basic rights that protect an individual's sphere of private autonomy, as well as a set of formally codified rules and procedures that both reflect social expectations regarding behaviour and conduct and facilitate the co-ordination of social action. The 'facticity' of these rules, rights and procedures refers to the fact that they are accompanied by enforceable sanctions that are to be levied in the case of transgression against them. It is the *fact* that violation of the law or the violation of another's circumscribed area of personal

autonomy will result in some form of punishment that encourages citizens or groups within a particular state to obey that state's laws.

If the facticity of modern law, that sanctions alone command compliance, were the exclusive reason why the rules and regulations set down by the state demand the observance of its citizenry, this would be called a strictly positivist account of the law. In contrast, Habermas contends that there is a complementary, co-original, and equally important dimension to law that is overlooked by strictly positivist explanations. This second aspect of the dual character of law is its *validity*, and this is the idea that citizens obey laws not only because they want to avoid the sanctions that accompany transgression but also because they believe there are good reasons why they *ought* to obey the law. Put another way, as citizens of democratic societies we do not comply with laws simply to avoid sanctions; we also do so because we view those laws as legitimate, that is, as deserving the respect of those subject to them. Validity, also construed as legitimacy, represents the normative dimension of law, a dimension captured in a principle that states "only those laws can claim legitimate validity if they meet with the agreement of all legal consociates in a discursive law-making procedure that has in turn been legally constituted" (Habermas, Facts and Norms 110).

The facticity dimension of law has a sociological basis insofar as Habermas addresses it from a "functionalist" perspective. The 'function' of modern law is to coordinate (via an elaborate structure of basic rights, statutes and codes) the social integration and interactions of those subject to the law by virtue of membership in a particular society and, in so doing, is the principle resource for a socially stable environment:

In functionally-differentiated societies, law has the special task of generalizing behavioral expectations in the temporal, social, and substantive dimensions in a congruent manner, such that contingently arising conflicts can be decided in a binding manner according to the binary code of legal/illegal.” (48)

From a normative perspective, however, modern law enjoins a philosophical discourse with our attitudes toward ideals of justice and rightness. In this view, law is not solely a coercive mechanism to bring about stable behavioral patterns but also encourages social stability insofar as laws and institutions reflect an interest in creating a *fair* and *just* society. Habermas, paraphrasing Rawls, agrees that a system of laws and basic rights can be said to have a “moral content” insofar as they embody the ideal that:

Just institutions would create circumstances under which it would lie in each one’s well-considered interest to pursue one’s own freely chosen life plan’s under the same conditions that also allow other persons to pursue their life plans. (58)

However, more importantly than expressing what one thinks just and fair institutions might look like is the need to describe “how these [institutions] can be *established* or promoted in present circumstances” (58).

The drawback of both the sociological and philosophical interpretation and explanation of modern law is that, when considered in isolation, each perspective ignores certain illuminating insights of the other, which in turn limits the explanatory power of each to show how modern legal institutions work. On the one hand, strictly normative interpretations fail because they cannot “bridge the chasm” between normative ideals and social facts. Narrowly moral explanations of modern law cannot adequately grasp the empirical constraints that, if not repudiate, then frustrate the conditions in which their ideals can obtain. On the other hand, systems theory conceptualizes law as an *autonomous* system of action that both “self-referentially delimits itself” from all outside communication and input and “is uncoupled from all other action systems” (49). The

problem with this perception is that it diminishes the capacity of law to have any *socially integrative* function. That is, if we accept the system-functionalist reading, “legal norms and legal acts lose all connection with the supposition of rationally motivated processes of reaching understanding within an association of legal consociates, [and] the validity claims and reasons put forth in legal discourses thereby forfeit their intrinsic value” (50).

If, as the means by which societies organize their affairs and interactions, modern law supposes a normative dimension (an orientation toward justice and fairness), then where in a complex society does this normative stimulus come from and how should we understand the way in which that normative quality can be reconciled with sociological explanations of law? Habermas begins this discussion by cautioning against too closely linking positive law with morality. This is because, while both legal systems and moral codes “serve to regulate interpersonal conflicts, and both are supposed to protect the autonomy of all participants and affected persons equally” (Habermas, Facts and Norms 450), positive law is distinct from morality insofar as its regulations are more concrete and less universal in their application, apply only to external coercible behaviour, and often involve questions that are not distinctively moral. Furthermore, the value pluralism that marks liberal democratic societies ensures an absence of a common, unified and substantive moral framework from which more concrete laws and regulations can derive their justification. Accordingly, if legitimate law has no “intrinsic moral content,” then any plausible normative interpretation of positive law must leave open two non-moral moments of self-determination: “the free choice of rationally deciding actors as well as the existential choice of ethically deciding persons” (451).

The issue, then, is how modern law maintains its legitimacy in light of these two competing, yet seemingly irreducible functions? That is, how do laws simultaneously encourage compliance amongst all subject to them, but leave open to the self-deciding individual the reasons why they choose to comply? According to Habermas, under the conditions of post-metaphysical justification, legitimacy is rooted within the connection between a system of law and the democratic process, as “the source of all legitimacy lies in the democratic lawmaking process” (89). Expanding on Habermas’s idea, Geert Munnichs writes that “the legitimacy of the [legal] order does not follow from its legality, but is dependent on the way in which the legal norms are produced by the political system of decision making” (Munnichs 185). Of course, there is no unified normative model for how democracies are supposed function. Typically, this disaccord is expressed in the traditional disagreement between “liberal” and “republican” models of democracy, and is best understood in terms of the conceptual priority each allots to the private autonomy of the individual and the collective expression of public autonomy, respectively. By way of a detailed and contrast-driven analysis between liberal and republican models of democracy, albeit a strongly ideal-typical one, Habermas tries to show how private autonomy (individual rights) and public autonomy (popular sovereignty) “mutually presuppose” and “reciprocally influence” each other in a way that can clarify how—without one taking normative primacy over the other—both “democracy and human rights form the universalist core of constitutional democracy” (Habermas, “Popular Sovereignty” 465).

Liberal models of democracy generally prioritize individual basic rights, which protect a sacrosanct domain of private autonomy, over popular sovereignty. On the liberal

view, democratic politics is no more than a collective expression of private wills that come together in an aggregate form to pressure the government to enact policies that favor their private spheres of interest. On the other hand, republican models of democracy prioritize political rights and characterize them in a positive sense:

[Political rights] do not guarantee freedom from compulsion, but guarantee instead the possibility of participating in a common practice, through which citizens can make themselves into what they want to be—politically responsible subjects of a community of free and equal citizens.... [Politics is an] opportunity in which the members of quasi-natural solidary communities become aware of their dependence on one another and, acting with full deliberation as citizens, further shape and develop existing relations of reciprocal recognition into an association of free and equal consociates under law. (Habermas, “Three Normative Models” 240-41)

As Habermas explains it, the liberal view fails because a pre-political set of basic individual rights in and of themselves cannot show “the source from which enacted law may draw its legitimacy” (Habermas, Facts and Norms 89). This impediment becomes apparent in cases where judicial bodies are required to mediate between basic rights that come into conflict with each other. Similarly, under conditions of increasing value pluralism, the republican explanation of democracy also falls short because its views assume, falsely, that there exists “the substantive support of a culturally established background consensus shared by the citizenry” (296).

Because of these limitations in both perspectives, Habermas proposes a ‘third way’ by which to bridge the theoretical gap between republican and liberal theories of democracy, a way that entails an explanation of how individual rights and popular sovereignty are not in competition with each other but must instead “mutually presuppose” each other. According to Habermas, democracies need both a framework of basic individual rights so as to create the important universal normative conditions of

fairness and openness, as well as institutional opportunities for expressions of popular sovereignty, whereby private individuals can come together to collectively and democratically “reach an understanding” about, and hence offer a legitimate grounding of (a justification for), the rules that govern their joint lives together. To some degree, any constitutional democracy with a system of basic rights and regulations already upholds these two ideas, since:

Both of these elements are already interwoven in the concept of positive and coercive law: there can be no law at all without actionable subjective liberties that guarantee the private autonomy of individual legal subjects, and no legitimate law without collective democratic lawmaking by citizens who, as free and equal, are entitled to participate in this process. (Habermas, “Reconciliation Through” 72)

For Habermas, what morality, law and democracy have in common is that they all express, in different ways, “the same *idea of self-legislation*” (Facts and Norms 90). However, cultivating a critical model based on this idea, a basis on which to judge already existing democratic arrangements, still requires a more comprehensive understanding of how all three expressions of autonomy are interconnected with each other.

1.2

The remainder of the first half of Habermas’s analysis is primarily intended to elucidate the connection between legitimate law-making, democratic procedures, and the conditions of normative justification suitable to expressing “the concurring will or the rational consensus of all participants.” Of central importance to this discussion is the concept of ‘postmetaphysical’ justification. At the core of this concept is the view that the conditions of modernity, marked by an increasingly distanced and reflexive approach to

inherited systems of belief and traditional practices, have made it gradually more implausible to call on a shared religious or cultural worldview as a source to justify the arrangements by which diverse and heterogeneous peoples live their lives together. Accordingly, in the absence of such a shared worldview, Habermas asserts that “practical orientations can in the final analysis be gained only from rational discourse, that is, from the reflexive forms of communicative action itself” (98).

These practical orientations take on a more tangible form as the system of laws that govern the common life of diverse individuals and groups. However, in the absence of any overarching moral authority, if the laws that govern a society are to have any claim to legitimacy, then the justification of those laws must be such as to convince those that exist within such a legal framework—people of disparate worldviews and value orientations—that those laws have a universal content. As Habermas notes, the “reasons that are convenient for the legitimation of law must, on pain of cognitive dissonances, harmonize with the moral principles of universal justice and solidarity” (99). In this context, the idea that considerations of justice must be aligned in some way with the democratic procedures, procedures charged with the responsibility for creating legitimate law, serves as the starting point from which Habermas uses his own version of moral argumentation, that is, discourse theory, to mediate between normative and empirical approaches to democracy. Discourse theory, or discourse ethics, is “a theory of practical reason that takes the form of a theory of argumentation” and aims “to recover moral objectivity in a posttraditional world no longer able to look to an overarching moral authority agreeable to all” (Rehg, Insight and Solidarity 23). For Habermas, all intersubjective language use has a built-in orientation toward reaching agreement or

mutual understanding. Assessing if one does have good reasons to accept a norm “can become clear only under the *pragmatic* conditions of rational discourses in which the only thing that counts is the compelling force of the better argument based on the relevant information” (103). Discourse ethics specifies, in the form of a discourse principle (D), that:

Only those norms can claim to be valid that meet (or could meet) with the approval of all affected in their capacity *as participants in a practical discourse*. (Habermas, “Discourse Ethics” 66)

In other words, the discourse principle “is supposed to tell us what sort of procedure is required to test for the validity of a proposed substantive norm” (Abizadeh 199).

The justification of moral norms, in particular, is a demanding standard of argumentation because it presupposes a *universal* inclusiveness in the rational answers that such discourses reach. For Habermas, to say that something is moral is to make “the claim that a norm lies equally in the interest of all,” which is to say “that the interests of *each person* be given *equal consideration*” (Facts and Norms 103). However, “given the improbability of [open and inclusive] communication, moral discourses are, as a rule, carried out in an advocacy fashion” (183). How, then, can this demanding standard of normative justification be incorporated, even if only in an ‘advocatory fashion,’ into a theory of democratic legitimacy and the legitimate generation of law? In answering this question, Habermas first observes that the majority of the “problematic validity claims” democratic societies wrestle with involve matters, what he calls ‘pragmatic’ or ‘ethical’ matters, whose reference system is not “humanity or a presupposed world community” (as it is for moral questions regarding justice), but specific ethical-political communities that have distinctive self-understandings and traditions. As such, “a legal community’s political will, which of course should accord with moral insights, also expresses an

intersubjectively shared form of life, existing interest positions, and pragmatically chosen ends” (152).

He warns, however, that recognizing the distinction between moral and legal questions is not a submission to the “ingrained prejudice suggesting that morality pertains only to social relationships for which one is personally responsible, whereas law and political justice extend to institutionally mediated spheres of interaction” (183).

Accepting this disclaimer, however, still does not explain how a standard of moral justification, which involves interpersonal relationships, can become operationalized as a democratic principle, which deals more with abstract institutional (as opposed to interpersonal) relationships. Habermas further distinguishes between the validity of norms and the democratic legitimacy of statutes when he writes:

The principle of democracy should establish a principle of legitimate lawmaking...[where] only those statutes may claim legitimacy that can meet with the assent of all citizens in a discursive process of legislation that in turn been legally constituted.Whereas the moral principle functions as a rule of argumentation for deciding moral questions rationally, the principle of democracy already presupposes the possibility of valid moral judgments.On the premise that rational political opinion- and will-formation is at all possible, the principle of democracy only tells us how it can be institutionalized, namely, through a system of rights that secures for each person an equal participation in a process of legislation. (110)

Pertaining to the theory of argumentation, this moral principle describes how rules of argumentation are “*internally* constituted” whereas this democratic principle refers to the how the rules of argumentation are “*externally* institutionalized.” As a theory of political legitimacy, discourse theory lays down conditions to normatively evaluate the fair and reasoned procedures only under which political authority (which in complex societies is the sole authority to create and enforce law) can be said to have democratic legitimacy.

Yet, because the republican notion that connects the exercise of political power with an expression of a unified will on the part of a collectively-acting citizenry has been already precluded as a satisfactory explanation of the political system, the association between concrete manifestations of political power and the consent of the citizenry must be expressed differently. Here Habermas establishes a distinction between *communicatively-employed power* and *administratively-employed power*, that is, between the political power derived from “the discursive formation of a common will” and the power to undertake “the implementation of the laws issuing therefrom,” respectively (150).

These two kinds of political power are connected through the medium of law, a connection based on Habermas’s observation that politics and law are ‘internally’ related and that political power should not be conceived as “externally juxtaposed” to law; rather, it is “*presupposed* by law and itself established in the form of law” (134). This relationship is explained as follows:

The law presents itself as a system of rights only as long as we consider it in terms of its specific function of stabilizing behavioral expectations. These rights can take effect and be enforced only by organizations that make collectively binding decisions. Conversely, these decisions owe their collective bindingness to the legal form in which they are clad. (133)

This relationship between law and exercises of political power is not a “*self-contained* exchange” since law can have “a legitimating force only so long as it can function as a resource of justice” (145). A society’s sense of justice, shaped by its recurrently changing awareness and understanding of particular practices as just or unjust, has its roots in a public’s perceptions and the interpretations of a public that subsists within those practices. These perceptions and interpretations are articulated in democratic processes, a set of practices central to the ongoing production of legitimate law if we consider that “at

the posttraditional of justification... the only law that counts as legitimate is one that could be rationally accepted by all citizens in a discursive process of opinion- and will-formation” (135).

The distinction, but also the connection, between opinion- and will-formation runs parallel to the division Habermas draws—significantly—between communicative and administrative power. For Habermas, communicative power represents a *motivating* force for the exercise of political power and is generated when free agreement obtains under the open and inclusive conditions of discursive intersubjectivity. In politics, this “consensus-achieving force of a communication aimed at reaching understanding” is located in public bodies and acts as an “*authorizing* force” on deployments by the state apparatus of coercive and collectively-binding, that is, administrative, political power. Law serves as the *medium* that joins these communicative and administrative power together in two ways; first, law expresses, through a set of basic individual rights, the conditions of mutual recognition that make communicative freedom possible; and second, in the form of the constitutional state, law is the “framework of statutory authorization” that gives legitimacy to exercises of administrative power within the political system, thereby fastening it to communicative power.

1.3

The second half of Between Fact and Norms examines how the normative idealizations of the discourse-theoretic conception of democracy can, given certain constraints posited by the social sciences, be incorporated into a more substantive institutional model of democratic politics. As James Bohman puts it:

The basic idea is to foster processes of communication and to design institutional procedures that at least make it more likely that political decisions will be based on reasons that would correspond to those that would emerge from a discourse under ideal conditions. (Public 181)

But prior to giving a sketch of his own model of deliberative democracy, Habermas pursues a line of criticism against both narrowly normative and empiricist models of democracy. This critique generates descriptive parameters that he eventually incorporates into his own model.

In outlining the central line of argumentation his analysis will pursue against non-normative explanations of democratic processes, Habermas writes:

A reconstructive sociology of democracy must choose its basic concepts in such a way that it can identify particles and fragments of an “existing reason” already incorporated in political practices, however distorted these may be ... [This approach] is premised simply on the idea that one cannot adequately describe the operation of a constitutionally organized political system, even at an empirical level, without referring to the validity dimension of law and the legitimating force of the democratic genesis of law. (Facts and Norms 287-88)

Failing to identify an ‘existing reason’ in current democratic practices, regardless of how distorted that reason may appear in those practices, is what Habermas argues leaves empirically descriptive ‘realist’ models of democracy unable to surmise the inescapable function democratic procedures and institutions serve in linking up political power with the rule of law. In keeping with this criticism, realist models of democracy tend to advance the view that the conceptual relationship between legitimate law and political power is measured by “the stability of a generally held belief,” by any license or through any means necessary, “in the government’s legitimacy” (289). From this perspective, legitimacy loses all connection with rational acceptability and is reduced to a byproduct of the superiority of some interests over those of others. Setting his own position against that of Werner Becker, Habermas argues that ‘realist’ theories of democracy like

Becker's are typically "cryptonormative;" that is, they presuppose a normative structure to make their accounts cogent, but that any description of this normative structure is curiously absent from their line of reasoning. An empiricist model of democracy that "purges the concept of power of just that normative authority it gains through its internal connection with legitimate law" finds its explanation wanting (289). For example, Habermas sees empiricist accounts, in explaining how existing political arrangements endure, as by and large emphasizing the perspective of an *observer*. These descriptions say very little about whether or how this perspective is reconcilable with the perspective of a *participant* in the political process. From the observer perspective, legitimacy is exclusively equated with stability—"stability" meaning a citizenry's consent to being governed—and it is typically expressed in a wholly *subjective* fashion through secret balloting mechanisms. If political legitimacy resides strictly in the stability defined by majority rule, Habermas asserts, this explanation cannot satisfactorily clarify "why norms passed by the majority should be accepted by the outvoted minority" (292).

Habermas contends that in order to explain why a persistently outvoted minority would continue to cooperate with a political order that rules in a way that consistently aligns with interests that the minority perceives not to be its own, empirical realists fall back on the notion that democracy, like any political form, is nothing more than a temporally limited "domesticated struggle for power" and that "democracy means nothing other than that the [numerically and/or symbolically stronger] part of the people rules the other for a set time" (292). The minority accepts things the way they are as a mere *modus vivendi* until such time as they gain an advantage that allows them to impose their norms on others who do not share the point of view that grounds those norms. But,

this does not explain why the majority, once firmly entrenched in a position of authority, does not simply exercise its power to steadily suppress the minority in a way that suits the continued pursuit of its interests. As all appeals to basic rights lose their normative character in such a picture, the empirical realist attempts to assuage any concern for a tyrannical majority by saying that such an event is precluded “by the majority’s fear of becoming a minority itself” (293). The ‘fear’ of the majority and the ‘prospects’ for the minority becomes the motivation “to observe the established rules of the game” by adhering to the basic freedoms and liberties guaranteed as constitutional rights. Under this interpretation, both the rulers and the ruled hedge their bets when it comes to power, be it gradually slipping away or slowly moving within reach.

Explanations like these—ones that reduce the multilevel and complex workings of constitutional democracy to a contestation of undisguised power between dominant interest groups—cannot explain from the ‘participant’ perspective how political decisions are *rationaly grounded*. If undisguised power plays really were the sole organizing principle of society, then why should the majority in any society, which holds the tangible power, agree to organize itself in a democratic fashion at all? Why would citizens in the minority, upon perceiving the hollowness of the supposedly democratic institutions they live under, ever go along with such a scheme? And finally, empirical realism is at a loss to rationalize why democracy should be preferred to any other system of governance and does nothing to diminish the potential for Hobbes’s benevolent despot, who could enforce the terms of our social arrangements without any need for the democratic authorization of the public, to come out of the background. If, as realist interpretations cynically depict it, all validity claims are of a strategic kind oriented

toward not communicative but purposive rationality, then there exists no procedural framework by which to adjudicate competing normative claims thorough argumentation and debate. Endorsement of such a conclusion problematically implies that nothing of political significance could be considered rationally decidable, making politics (and law) as a whole an irrational enterprise.

Habermas is equally critical of normative models of democracy and what he perceives as their failure to fully grasp certain social facts (particularly those elaborated in systems-functionalist theories) that either impede or erode their normative aspirations. Functionalist explanations note how capitalist modernization and accompanying increases in the size and complexity of the administrative welfare state have given rise to the market economy and an organized bureaucracy. As the administration and economy grow in size and sophistication, they are increasingly detached from a rationality engraved in the “communicative practice of everyday life,” the ordinary communication that makes up what Habermas calls the “lifeworld.” This escalating detachment from the background culture of everyday traditions, practices and values makes these systems increasingly, if not insuperably, resistant to democratic organization. While these self-regulating economic and administrative “systems” do “represent a higher and evolutionarily advantageous level of integration by comparison to traditional societies” (Habermas, Theory of Communicative Action 339), in that they have exponentially increased the net wealth and efficient organization of society, the administrative-economic complex increasingly enforces its own purposive rationalities and goals into other areas of previously distinctive traditional and cultural life. In this picture, what started out as a societal rationalization, one that harnesses and mobilizes the material and

cognitive resources of a society, becomes a *colonization of the lifeworld*, where the systemic imperatives of money and power tend to restrict the spheres of action where individuals can use ordinary language to co-ordinate their lives.

Habermas considers this aspect of modernization, the emergence of autonomous spheres of action that are self-regulating, to be a largely irreversible trend that any proceduralist model of democracy must be attentive to. An increased appreciation of the functional differentiation of society is what Habermas believes makes his democratic proceduralism more plausible than Joshua Cohen's version. Habermas maintains that Cohen's proceduralism is dogged by "the idea of a society that is deliberately structured *as a whole*," and that tying democratic legitimacy to a procedure will work "as the core structure in a separate, constitutionally organized political system, but not as model for all social institutions" (Facts and Norms 305). The alternative is to view "the political system as neither the peak nor the center, nor even as the structuring model of society, but as just *one* action system among others" (Habermas, "Three Normative Models" 251).

However, authorized political bodies are unique action systems in that they have sole control over the establishment of legal regulations and statutes, and law is not just an action system like any other. The legal system is unique because it has both *systemically* and *socially* integrative function:

From the standpoint of *social theory*, law fulfills socially integrative functions; together with the constitutionally organized political system, law provides a safety net for failures to achieve social integration. It functions as a kind of "transmission belt" that picks up structures of mutual recognition that are familiar from face-to-face interactions and transmits these, in an abstract but binding form, to the anonymous, systemically mediated interactions among strangers. (Facts and Norms 448)

Furthermore, law also “functions as a hinge between system and lifeworld,” one that “transforms” and “translates” the ordinary language of everyday communications into the language of law, putting ordinary language “into a form that is comprehensible to the special codes of the power-steered administration and the money-steered economy” (56 & 81). Thus, politics is able to “communicate through the medium of law with all the other legitimately ordered spheres of action, however these happen to be structured and steered” (302).

The central concepts that comprise Habermas’s analysis—the way private and political autonomy presuppose each other, the legitimacy of laws and exercises of political power rely as much on their justifiability to those that are subject to them as on guarantees to their enforcement, the relationship between legitimate law and postmetaphysical principles of moral justification, and the relationship between the motivating force of discursively-generated communicative power and the binding force of administrative power—lay the groundwork for his own more substantive program of deliberative politics and democratic proceduralism. Habermas begins by characterizing the nature of the procedures that make up his program:

The democratic process is governed by *universal* principles of justice that are equally constitutive for every body of citizens. In short the ideal procedure of deliberation and decision making presupposes as its bearer and association that agrees to regulate the conditions of its common life *impartially*. (306)

From this, we can see that Habermas places strong emphasis on the inclusiveness and *neutrality* of democratic procedures. Borrowing from Cohen, Habermas outlines several guiding principles that, through their approximation in practice, equip democratic procedures with their claim to legitimacy. These postulates include: that political deliberation is argumentative in form and involves the critical exchange of reasons; that

political deliberation is public and inclusive and gives everyone an equal opportunity to be heard, both in terms of proposing ideas, putting forward information, and critically analyzing the proposals of others; that political deliberators are bound only by “the presuppositions of communication and rules of argumentation;” and that no coercion, internal or external, is permissible that might interfere with a deliberating participant’s ability to criticize or make proposals (305). Collectively, these principles outline a standard of legitimacy, to be expressed in the form of an ‘ideal’ procedure, by which citizens jointly determine the rules of their common life together.

However, democratic governments must take decisions within a limited amount of time, and cannot deliberate indefinitely on the particular course of action they will choose to pursue. As such, decisions made at the conclusion of deliberative processes will not directly be the result of an idealized deliberation, but are almost without exception brought through a majority vote. The question, then, is how to link-up this non-deliberative mechanism of reaching decision with the deliberative activities that precede that decision. Here is where the affiliation between communicative and administrative power takes on institutional significance, and the distinction between the two is recast by Habermas as one between “arranged” publics (the public sphere) and “weak” or “informal” publics (parliaments or other empowered political bodies). Administrative systems rationalize the decisions they take based “not on the practical reason involved in applying norms but the effectiveness of implementing a given program” (Habermas, “Popular Sovereignty” 483). The instrumental considerations of the administration can only be normatively ‘steered’ by parliamentary bodies. Parliaments are “arranged” public bodies, distinguishable from more general public bodies insofar as they are “socially

bounded and temporally limited” and, more importantly, have the capacity to make binding decisions. These parliamentary bodies are “structured predominantly as a *context of justification*,” and they deal primarily with problems of social organization by way of “cooperative solution of practical questions, including the negotiation of fair compromises” (Habermas, Facts and Norms 307).

“Weak” publics, by contrast, do not refer to a specific body or organized institution but to a ‘political public sphere,’ defined as an “open and inclusive network of overlapping, subcultural publics having fluid temporal, social and substantive boundaries” (307). The ‘informal’ nature of the public sphere refers to the more or less “spontaneous” and “subjectless” associations of which it is comprised, associations in which all individuals are guaranteed the same set of constitutionally-protected basic rights. Keeping the public sphere at some remove from administrative political power serves to “*relieve the public of the burden of decision-making*,” a feature that Habermas asserts has an “*intellectualizing effect*” on the communicative interactions that take place there. By detaching it from the “concrete practical obligations” that follow from institutionalized democratic procedures (the ‘obligation’ to make timely, decisive and binding political choices), the “anarchic structure” of the public sphere may leave it organizationally unequipped to generate specific solutions to problems in the way that formal parliamentary bodies do, although there is nothing to suggest they are incapable of doing so. More importantly, this independence from the pressures of having to make binding decisions heightens its potential to act as a “sensor” to detect society-wide problems that more formal bodies are insensitive to. In the informal public sphere,

New problem situations can be perceived more sensitively, discourses aimed at achieving self-understanding can be conducted more widely and

expressively, collective identities and need interpretations can be articulated with fewer compulsions than is the case in procedurally regulated public spheres. (308)

Of course, the definitions of both 'weak' and 'arranged' publics, is somewhat abstract and highly idealized, thus making them vulnerable to sociological disenchantment. This is why the distinction Habermas draws between the two publics is, as we shall see in the next chapter, frequently brought up by his critics. For instance, these idealizations of communication, while sketching out indispensable elements for a rational political will-formation, leave open lots of questions about the attitudes, motives, and competences of participants. It also says little about structural inequalities that create disproportionate, disparate opportunities to access and control information. Moreover, decision-making in democratic societies is subject to countless pressures that tax the capacities of the bodies responsible for those decisions to exercise their political power in a way that is inclusive of the reasoned opinions of the citizenry.

However, in looking to overcome these sorts of problems, Habermas insists that the communicative power that originates in the political public sphere "can be converted into political power only if it passes through the sluices of democratic procedure and penetrates the constitutionally organized political system in general" (327). Simultaneously, any analysis of this "*official*" circulation of power must also concern itself with "how effectively the *unofficial* circulation of power encroaches on the constitutionally regulated circulation of power" (328).

Habermas begins this analysis by remarking that a "sociological translation of the discourse theory implies that binding decisions, to be legitimate, must be steered by communication flows that start at the periphery and pass through the sluices of democratic and constitutional procedures situated at the entrance of the parliamentary

complex” (356). In this ‘core/periphery’ model, the legitimate “official” circulation of political power, which necessarily aligns with constitutional principles, is viewed in terms of how communicative ‘influence’ generated in the public sphere acts on parliamentary bodies and the judicial system near the core, and ultimately affects the ‘implementation’ of actual policy by the administrative complex.

At the same time, there is “countercirculating” power competing with the official channel. To take an example, governments in large countries necessarily rely on sizeable administrative bodies with specialized functions to deal with the mundane but indispensable details, such issues as allocation and distribution, which must be attended to in implementing the, by comparison, relatively vague directives of parliament. Conflicts between what parliaments want and what administrative bodies can provide signifies a kind of administrative power completely cut off from the legitimating influence of communicative power. Nonetheless, the relationship between modern governments and various administrative bodies follows an “established pattern,” and importantly relieves elected governments of the “unavoidable complexity” that arises from their decisions; conversely, this same relationship diminishes the potential of the public sphere to directly exert influence over decision-making bodies at the core. The solution to this inescapable countercirculation of power, Habermas asserts, is to look at how this insoluble arrangement can still “remain open to renovative impulses from the periphery” (357). His own proposal is to alter the “mode of operation” assumed by the political public sphere, changing that mode from one of straightforward, direct influence on political bodies (the power to directly bring about specific decisions) to one of

problematization. In terms of this function, Habermas defines an effective public sphere, given its limited capacities to directly 'solve' problems, as:

A warning system with sensors that, though unspecialized, are sensitive throughout society. From the perspective of democratic theory, the public sphere must, in addition, amplify the pressure of problems, that is, not only detect and identify problems but also convincingly and *influentially* thematize them, furnish them with possible solutions, and dramatize them in such a way that they are taken up and dealt with by parliamentary complexes. (359)

With this definition of the public sphere, communicative power takes on an "informal" and indirect configuration in contrast to formal and binding exercises of political power. The public sphere remains tied to tangible exercises of political power only insofar as the legitimacy of that power is measured in terms of how effectively the voices and concerns of the public influence the deliberations of democratically-elected representatives who enact binding legislation and policy. Why does Habermas only assign the public sphere this informal, yet indispensable purpose in the democratic process? It is because he thinks the quality of political deliberations in the public sphere "is dependent on the extent to which the process satisfies the communicative requirements for a free and unconstrained discourse, in order that all relevant arguments will be heard" (Munnichs 186). By confining actual decision-making to the administrative apparatus and parliamentary bodies, public spheres are relieved from the distorting effects of temporal-spatial constraints which, as a matter of efficacy and efficiency, must necessarily be placed on deliberation.

Chapter Two: Deliberation, Discourse and Consensus

To continue developing from the thematic materials surveyed in chapter one, especially the distinction between the informal public sphere and formal decision-making bodies, Habermas's political theory invites two lines of critical objection that I now draw attention to. The first line of criticism focuses on Habermas's effort to "transform the public sphere into an arena of critical autonomous debate that is insulated from the distorting effect of power and money" (Chambers, "Politics of Critical Theory" 231). What Habermas considers a virtue in trying to bridge normative and sociological perspective on democracy—relieving the public sphere of actual decision-making power so as to approximate the conditions required for intersubjectively reaching understanding, thereby confining its democratic potential to the *perception* of problems and the *thematization* of influential opinions rather than giving rise to actual binding *solutions*—his critics see as either an "unwillingness to engage in a genuinely emancipatory politics" (235) or an acquiescence to "politics as usual." One promising avenue of inquiry as to why Habermas's project is not, according to some critics, more democratically ambitious in scope is the "specific categorical bifurcation" he makes between systems and the lifeworld. I agree with Douglas Kellner's assessment that Habermas "makes too rigid a categorical distinction between system and lifeworld, constructing each according to their own imperatives, thus removing the "system" (that is, economy and state) from democratic transformation, while limiting the site of participatory democracy to the lifeworld" (274). In fact, it is in the context of this lifeworld/systems distinction that this thesis first considers two 'political-theoretical' criticisms that take exception to what they see as Habermas's insufficiently imaginative democratic politics. Against this

interpretation, and in light of a second line of criticism directed at Habermas, I will argue that these same issues can be more adequately explained as the product of incorporating the strong idealizations of discourse theory into a material political theory of social arrangements and institutions.

This second line of critique engages Habermas's political theory in terms of the theory of moral justification upon which it is modeled. These analyses consider the inexactitude with which Habermas couples the communicative power of the public sphere with the formal decision-making powers of centralized governing bodies as a problem symptomatic of an unresolved tension between the strong justificatory idealizations of discourse theory and the ambiguities of deliberative democratic politics. These critics take as their point of departure the demanding standard of moral justification suggested by Habermas's discourse ethics and the modifications it undergoes in its transformation into a normative theory of democratic legitimacy. Their assessments conclude that certain guiding idealizations of a practical discourse, notably discourse ethics' emphasis on (1) the capacity of discursively acting agents to impartially assess and abstract themselves from their particular value commitments; and (2) the stipulation that any question capable of being assessed rationally also admits of ideal rational acceptability (that is, the question involved is capable of commanding unanimous agreement for identical reasons), are too demanding a standard for political agreement and are, therefore, an unsuitable criterion upon which to base a model of institutionalized deliberative politics. More importantly, from the perspective of this thesis, Habermas's commitment to bringing these demanding idealizations as close as possible into democratic praxis inadvertently cause him to overlook other potential means for

democratic decision-making and conflict resolution, specifically decision mechanisms involving compromise formation and negotiations in which the public sphere can play a far more active role than Habermas ascribes to it.

While both lines of critique expose prospective weaknesses with Habermas's program, they by no means prove that a procedural model of democratic politics based in discourse ethics is irretrievably flawed. In the section that closes out this chapter, I will try to provide evidence that the root of the problems identified in these critiques, especially certain features of Habermas's discursive idealizations, are more a problem with his moral theory than a failure of democratic imagination. This explanation may go some way to helping explain the deep doubts some authors express toward certain institutional features of his more substantive political program. If accepted, this opens the possibility for suggestions as to how a modified interpretation of the discourse theory of democracy, in particular one that makes more allowance for compromise formation and alternative modes of conflict resolution in political will-formation, can be more responsive to some of the criticisms leveled in this chapter.

2.1

"Modern societies are integrated not only socially through values, norms, and mutual understanding," Habermas observes, "but also systemically through markets and the administrative use of power" (Facts and Norms 39). Social integration takes place in the *lifeworld*, a background of shared meaning that is the sum of all communicatively structured interactions. The lifeworld serves as the setting against which agreement is reached (and disagreement is tested) about the common rules, norms and values that give

order to a society's collective life together. System integration, on the other hand, refers to the emergence and proliferation within these societies of market economies and administrative bureaucracies. Eventually, as societies become socially differentiated—that is, more specialized in terms of greater pluralization (increasingly diverse value horizons) and individualization (self-interest) within the population—the face-to-face establishment of valid norms by interpersonal communication is overwhelmed, and ongoing instances of explicitly interpersonal contact for reaching a common understanding are gradually replaced by a system of law that can mediate dissensions about shared understandings more anonymously. Concurrent with these increases in value pluralism and personal individualization, modern societies have also seen the rise of global capitalist markets and vast increases in the size and role of government that accompanied the birth of the welfare state. The basic idea of systems integration is that markets and organizational bureaucracies—having gradually grown in ubiquity and complexity—eventually begin to function autonomously, obeying only their own purposive logics of increased profitability and efficiency, and consequently resist any conscious efforts toward social control. Eventually, these “systems” start to steer society independently or “behind the backs” of those affected by them. Free markets and organizational bureaucracies are systemically integrative, rather than communicatively or socially integrative, because the legitimation of their activities relies not on an intersubjectively achieved mutual understanding amongst all those impacted by the effects of their activities but by the instrumental achievement of some set of non-communicatively determined ends (in this case the maximization of profit and the efficient implementation of government policies requiring the allocation of scarce

government resources). Thus, the administrative state and the capitalist economy become autonomous systems that have grown largely detached from the conscious control of society. And, while the administrative system does retain some connection to democratic governance in that its goals are primarily programmed by political representatives in the form of elected officials, in the case of the economy its “capitalist growth dynamics can be preserved only if ‘the propelling mechanism of the economic system [is] kept as free as possible from lifeworld restrictions as well as from demands for legitimation directed to the administrative system’” (McCarthy 162).

William Scheuerman argues that Habermas’s project is too easily inured to the resistance that markets and (to a lesser degree) bureaucracies apparently show (at least by way of a systems-functionalist analysis) to further democratization or social control. Moreover, by not offering anything in the way of a systematic analysis of the way capitalism and bureaucracies reproduce and exacerbate social asymmetries of power, *Between Facts and Norms* can offer only an “inadequately critical assessment” of existing democratic arrangements, thus leaving discursive proceduralism with at best a “tangential relationship” to capitalist democracy. Scheuerman supports his position by first indicting the “ambivalent” connection Habermas draws between the communicative power generated in the “anarchic” and “spontaneous” networks in the public sphere and the administrative power that generates actual public policy and binding laws. As discussed above, Habermas believes that opinion-formation in civil society can only exert a non-specific communicative “influence” on political will-formation. As there are no explicitly defined parameters or explanation of how communicative power is transformed into administrative power, the degree and extent of this influence is indeterminate.

Furthermore, since this influence has no binding effect of political will-formation, according to Scheuerman, the communicative potential of the public sphere is narrowly conceived as a “defensive” tool against administrative blindness to society-wide problems that go undetected, rather than a formidable tool by which to more directly rationalize an administration’s authority.

Scheuerman attributes this reduced capacity of public deliberation to make binding claims on political actions as the result of a curious “fundamental conceptual tension” in Habermas’s analysis of how to connect communicative power with administrative power. He situates this tension in Habermas’s odd co-mingling of two radically different perspectives on democracy, those of Nancy Fraser and Peters, into his own two-track political model.

The way Habermas draws on Fraser’s and Peters’ work is something of an elision, since he utilizes only fragments of either perspective to outline a concrete model of deliberative politics whose procedures resemble the neutrality and impartiality embedded in the demanding presuppositions of argumentative discourse. This undertaking requires adapting the normative idealizations of discourse and the “unrestricted communication community” to the empirical realities and cognitive limitations of public deliberation. So, for example, a rule of discourse whereby “everyone is allowed to introduce any assertion whatever into the discourse” (Habermas, “Discourse Ethics” 89), becomes the deliberative maxim that insists that “*any* topic that at least one participant considers to be publicly relevant must also be a valid item for public discussion” (Habermas, Facts and Norms 312). In other words, it cannot be decided prior to deliberation what is and is not suitable for public debate, for this question is a matter

that can only be determined through the carrying out of actual deliberations. However, Habermas stipulates that “every affair in need of political regulation should be publicly discussed, though not every legitimate object of public discussion will in fact be politically regulated” (313). Hence for Habermas it is the give and take of reason in civil society where problems are detected, thematized, and eventually determined to be or not be in need of intervention by the regulatory state. This perspective is congruent with Fraser’s view, a view which Habermas himself evokes, that “only participants themselves can decide what is and what is not of concern to them” (312).

But, as Scheuerman notes, Habermas seems to stray from this position when he adopts certain points of Peters’ core/periphery “sluice model” to describe how legitimate political power circulates through society. From this position, the balance of binding political power inevitably slides away from the periphery and centralizes in the core, and the “perception” and “thematization” of problems increasingly falls to parliament rather than civil society. Concomitantly, “now it is the administration,” and not parliament, “that is seen as possessing the most impressive capacity for handling and resolving problems.” (Scheuerman 72). By accepting without reservation Peters’ argument that in complex societies “the political center inevitably gains independence in relation to the periphery,” Scheuerman charges that Habermas relinquishes the potential for the further democratization of society to the “normal” politics of the everyday, and in so doing he fails to show “sufficient concern for [this trend’s] potentially worrisome implications for democratic politics” (73). This imprecise theoretical reconciliation of Fraser’s “democratic socialism” with Peters’ “cautious democratic realism” provokes doubt for Scheuerman as to whether Habermas convincingly connects the communicative power of

the informal public sphere with the administrative power that establishes binding decisions and policy. For, as Scheuerman notes, at times Habermas argues that:

Communicative power can rely on the medium of law to *determine* administrative power. Yet, at many other junctures, Habermas offers a more modest view of the scope of communicative power: communicative power “more or less”(!) programs, and merely “influences” and “countersteers” administrative power. (Scheuerman 70)

The reason why Habermas’s construction of the relationship between communicative and administrative power appears ‘ambivalent,’ I believe, is that the relationship between the two is not a static one. Any unambiguous correlation between the two is obviated by both the (irreversible) functional differentiation of society and the notion that communicative power “can develop only in undeformed public spheres; it can issue only from structures of undamaged intersubjectivity found in nondistorted communication” (Habermas, Facts and Norms 148). Nevertheless, as to what constitutes an ‘undeformed’ public sphere, and to what extent the functional differentiation of society inhibits democratic impulses are both matters that Habermas can be confronted on.

At the base of his entire critique lies Scheuerman’s more specific concern that, as a theory of deliberative democracy, *Between Facts and Norms* in the end fails to elucidate how “deliberative democratic ideals could be institutionalized within the very *core* of state bureaucracy” (78). This lacuna ultimately leaves Habermas without anything systematic to say on how social asymmetries undermine democratic aspirations. Furthermore, this limitation also casts doubt as to the efficacy of Habermas’s suggestion that evolving, autonomous “social movements and new subcultures” (Facts and Norms 382) present themselves as a means by which citizen input can act as a “reactive control” on the “fortress of the state apparatus” (440).

On the surface, Scheuerman's analysis resonates with Iris Marion Young's efforts to "sound a caution about trying to put ideals of deliberative democracy into practice in societies with structural inequalities" (Young 671). Young warns that such deliberative aspirations are misguided in that they overlook the "subtle" ways in which persistently unjust social arrangements and structural inequalities create a "discursive hegemony," one that manages the political agenda from the outset, thereby excluding challenges to enduring and unaltered premises that inform debate. The thrust of her argument is that long term inequalities between citizens in terms of wealth, education, and social access or privilege serve to "deeply influence the premises and terms of discourse that make it difficult to think critically about aspects of their social relations or alternative possibilities of institutionalization and action" (686). The absence of a "genealogy of discourses" by which to expose society's shared conditioned assumptions—assumptions that not only thematize what social trends count as political 'problems' but also narrow the possibilities for working those problems out—leads Young to conclude that participating in political deliberation under such circumstances can only serve as a consent to the further maintenance of social injustice. Contrary to Scheuerman, Young concludes by offering encouraging alternative examples, mainly in the resistant vein of civil disobedience and creative modes of political activism and expression, by which "persistently unequal citizens" can challenge the 'reasonableness' of existing democratic procedures and institutions under "existing structural circumstances." However, Young's suggestions go beyond the "acts on nonviolent, symbolic rule violation" Habermas sanctions only when governments issue "binding decisions that... [civil protestors] consider illegitimate in the light of valid constitutional

principles”(Habermas, Facts and Norms 383), for she feels that protest and disobedience is needed whenever “existing social and economic structures... set unacceptable constraints on the terms of deliberation and the agenda” (Young 682).²

It is important here to remark that Scheuerman and Young seem to assail the similar elements of Habermas’s democratic theory: namely, the separation of formal decision-making powers and agenda setting from the communication structures of the public sphere. Nonetheless, they end up reaching very different conclusions regarding ameliorative possibilities. Whereas Scheuerman thinks that making politics more deliberative and inclusive involves instilling certain ideals into the institutions where formal and binding power is exercised, Young thinks it more important to continuously and unrelentingly challenge deliberative pretensions, as they are masking latent coercive forces and repressive intentions. Nevertheless, both stances intimate that there is an as yet unresolved tension in Habermas’s work between politics as a systemically-integrated administrative activity and politics as a socially-integrated public activity.

In seeking to explain this tension, Thomas McCarthy makes the case that Habermas overuses the theoretical apparatus of systems theory to describe the rationalization of society, and in so doing he leaves the critical potential of communication theory to expose the deformities that have accompanied such rationalization in an “unnecessarily defensive position.” The problem with systems theory is that it takes, what McCarthy calls (echoing Habermas), a “counterenlightenment” perspective towards intersubjective rationality: “[Systems theory] represents...the most refined form of the technocratic consciousness [because it] serves to *insulate practical questions from public discussion*” (McCarthy 169). While

Habermas is highly critical of systems functionalism in his democratic theory, and he furnishes a convincing case against the systems-theoretical viewpoint that society can be systemically integrated as a whole,³ systems theory nevertheless continues to exert a great deal of influence on his account of social complexity. For example, in delimiting the potential for political activity to steer society as a whole, Habermas submits that:

Deliberative politics remains a component of a complex society.... [and] the political system remains dependent on other functional mechanisms, such as the revenue-production of the economic system, in more than a trivial sense.... deliberative politics, whether realized in the formal procedures of institutionalized opinion- and will-formation or only in the informal networks of the political public sphere, stands in an internal relation to the contexts of a rationalized lifeworld that meets it halfway. ("Three Normative Models" 251-52)

Although this statement is representative of Habermas's general effort to set apart his own democratic program from the 'sociological naivety' of Rawls or Joshua Cohen, each of whom he accuses of being insufficiently attentive to the social contexts and empirical constraints that provide obstacles that frustrate normative aspirations, such a statement also seems to limit *from the outset* the emancipatory normative potential of communicative rationality.

While I agree that a normative theory of democracy, to avoid being labeled utopian or impractical, must provide some account of how certain social processes seem to function as independent spheres of action (largely out of the direct control of individual or group consciousness), the degree to which these processes resist democratic steering is an *empirical* question for which Habermas is taken to task. For, there is the question of just how distinct the hierarchically-organized and impersonal administrative system is from representative parliamentary bodies, the answer to which provides a further sense of how closely aligned the latter is with the communicative influence of the

public sphere. In asking whether the political sphere is systematically or socially integrated, that is, does it deploy communicative or instrumental rationality in making its decisions, McCarthy concludes that the answer is “ambiguous at best” and that any “either/or” answer will be unpersuasive.

This is because a strong case has been made against viewing any organization, even an administrative one where there are plainly defined rules regarding patterns of behaviour and where power relationships between superiors and subordinates are clearly laid out, as strictly formally organized. Most formal bureaucracies, despite the way rigidly defined roles and objectives direct their activities, incur a great deal of informal organization whereby interpersonal values such as a shared orientation to norms and culture have an effect on the actions of the organization. Interactions within “functionally specified” and “media-steered” organizations like a complex administrative apparatus rely for the coordination of their activities not only on “functional interconnectedness” (for example, clearly defined formal roles and hierarchical power structures), but also on less formal open channels of communication (the necessity for justification and mutual understanding). This indicates that “there are situations in which organizational superiors can act authoritatively only “with reservation,” that is, in which they know they cannot achieve their goals without collegiality, cooperation, mutual understanding” (McCarthy 167). Extrapolating from these observed differentiations between communicative and purposive rationality even within systems themselves, one might conclude that the demarcation between system and lifeworld, particularly in the political sphere, is an uncertain matter that is highly contestable, leaving open the possibility for the further democratization of the parliamentary-administrative complex.

2.2

I now want to turn to a second set of criticisms that draw more specifically on the obstacles associated with basing a substantive political theory on the foundations of discourse ethics. Not only do these analyses, I want to argue, capture the dilemma that arises from applying the highly idealized presuppositions of rational argumentation to all of the recalcitrant vagueness and contingencies of actual social structures and political institutions, but they also have something of a complimentary relationship to the kinds of dissatisfaction that Young and Scheuerman express toward Habermas's project, especially how he distances the democratic influences of civil society from binding exercises of political power.

Habermas is not insensitive to the distinction between highly reflexive moral discourses and actual political deliberations between flesh and blood participants. In fact, the realization of this distinction is a central concern of his attempted reconciliation of normative and sociological perspectives on democratic politics. As a moral theory, discourse ethics "does not by itself generate substantive moral norms; rather, it specifies an argumentative procedure that any norm must satisfy if it is to be morally acceptable" (Baynes 109). When utilized as the basis for a framework by which to assess the legitimacy of democratic decision-making, discourse theory is used to evaluate the justness and fairness of institutional arrangements and procedures from which specific policies and decisions descend. Thus, one way to productively analyze this connection is to critically assess how successfully Habermas maintains this distinction between the standard of justification set out in his moral theory and application of this standard to his political model. As my assessment will show, when taken in the context of Habermas's

critics discussed below, in the adaptation of discourse ethics from a moral to a political theory certain previously identified deficiencies with the moral theory not only persist but are exacerbated. If this proves to be the case, the next step will be to ascertain the specific nature of these problems and ask if they can be dealt with in a way that effectively preserves the critical potential of discourse theory to further the democratic franchise. These alternatives will be explored at length in chapter three.

There is one issue in particular, albeit an issue with many different expressions, that is central to this discussion and is prominent in the second group of “discourse-oriented” critiques: Habermas’ insistence that rational argumentation must entail a ‘convergence toward consensus’ amongst participants (concerning the agreement that is ultimately reached). The emphasis on the consensus-producing force of reason, an idea that plays a central part in his discourse ethics, leads the authors I will discuss to conclude that Habermas makes too many strong assumptions about both the nature of the problems that feature in political debates as well as about the participants that are party to political deliberation. The refutation of these assumptions serves as the point of divergence from which other discourse models of deliberative politics are conceived. Prior to looking at these discourse-based evaluations of Between Facts and Norms, and showing how they more persuasively explain the same problems cited in the political-theoretical evaluations discussed in section one, it is necessary to briefly explain discourse ethics and why the feature of consensus plays such a prominent role in it.

As already mentioned, discourse ethics is a moral theory. It is built on Habermas’s response, developed out of his earlier analysis of speech acts and a reconstruction of the unavoidably rational structure of communication, to a moral

skepticism that argues moral judgments do not have any cognitive content and therefore cannot be rationally justified. Habermas argues that normative validity claims (that one 'ought' to behave in such-and-such a way) are not simply the expression of volitional inclination but do have a rational core that allows them to be rationally justifiable. The thrust of this claim is that we agree with normative statements not because they coincidentally happen to align with our own beliefs or desires but because we believe that, if called upon and under certain conditions, these statements could be discursively redeemed with good reasons. Discourse ethics is thus "an attempt to reconstruct basic moral intuitions already contained in our communicative practices" (Baynes 73) by "[reconstructing] the moral point of view in terms of rational discourse, that is, real practices of argumentation and dialogue that strive to meet idealized conditions of reasonableness" (Rehg, Insight and Solidarity 83). In turning a theory of communicative rationality into a theory of communicative ethics, Habermas takes the "complex relations of mutual recognition" embedded in communicative practices and gives them normative form in a definite set of 'rules of argumentation.' These 'rules' are essentially conditions that define a rational discourse as one that is inclusive of any competent speaker wanting to participate, that allows any speaker to be able to challenge any assertion tabled and introduce any assertion she wants to, and that takes place in the absence of any internal or external coercive forces. Together these conditions specify a procedure that is encapsulated in the form of a discourse principle (D), which states: "only those norms can claim to be valid that meet (or could meet) with the approval of all affected in their capacity *as participants in a practical discourse*" (Habermas, "Discourse Ethics" 66).

One objection typically registered against this interpretation of ‘morality as an argumentative practice’ is that norms are invariably couched in particular historical forms of life, perhaps cultural or religious. And, if norms actually are culture-specific, this limits their potential to be rationally debated across groups that do not share the same value horizon presupposed by the norm itself. Recognizing the strength of this objection, Habermas introduces a corollary principle that gives content to what he views as the appropriate *object domain* of moral theory, namely, whatever matters that can be said to lie in the equal interest of all. This reflects Habermas’s intuition that judgments of justice and fairness transcend local affiliations. This demarcation of the moral domain then assumes the form of a universalization principle (U) that states a norm is not valid “unless all affected can *freely* accept the consequences and side effects that the *general* observance of a controversial norm can be expected to have for the satisfaction of *each individual*” (93). For Habermas, (U) is more than simply an ideal or standard of inclusiveness that valid norms should strive to approximate; it is more like an “implicit” stipulation attached to recognition of those ‘inescapable’ rules of argumentation embedded in communicative rationality. As such, (U) encapsulates “what convincing or rationally motivating reason giving involves,” namely, that “*if* an acceptable norm emerges from discourse at all, it must embody in some way a general interest” (Rehg, Insight and Solidarity 39). The universalization principle acts then “like a knife that makes razor-sharp cuts between evaluative statements and strictly normative ones, between the good and the just” (Habermas, “Discourse Ethics” 93). Questions of justice are moral precisely because they involve *generalizable interests*, that is, everyone has the exact same interest in being treated justly or fairly, while ethical questions are embedded

in an individual or group's particular, and non-generalizable, version of the good life. Subsequently, while ethical problems "are accessible to rational discussion only *within* the unproblematic horizon of a concrete historical form of life or the conduct of an individual life," moral problems "can in principle be decided rationally" because only these kinds of problems are capable of *unanimous* consent, that is, they "hold out the prospect of *consensus*" (Habermas, "Discourse Ethics" 104).

As we can see, Habermas confines the object domain of his discourse-based moral theory to issues that are generalizable and therefore hold out the prospect of unanimous agreement, both of which align with a central theme in his moral non-cognitivism, namely, that matters of justice have ingrained in them an "intuition that is very hard to deny... the idea of impartiality" ("Discourse Ethics" 72). Critics of communicative ethics, notably those supportive of the endeavor on the whole, have pointed to weaknesses not only with the moral/ethical distinction Habermas constructs but also, and more importantly for the purposes of this paper, have raised concerns regarding the practical import of concepts central to discourse theory into political philosophy; in particular, the ostensibly 'intuitive' characteristics of impartiality, generalizability and unanimity connected to the moral point of view. These lines of inquiry ask both whether the uncompromising standards of moral discourse can convincingly reply to deliberative democracy's critics and their familiar calls of 'empty formalism' or 'utopianism,' and whether Habermas's own discourse theory of democracy might not inadvertently transgress his own monition that "one shouldn't confuse discourse with the legally institutionalized procedure of political will-formation" (Habermas, "Discourse Ethics, Law" 256). Regarding this last point, these criticisms ask

if the highly demanding regulative ideals of discourse theory might cause unforeseen problems if drawn on as guiding principles for political deliberation.

For example, James Bohman and William Rehg question what they call Habermas's "strongly epistemic interpretation" of deliberative democracy ("strong" in that it references discourse ethics' highly ideal standards of impartiality, unanimity and consensus) and the way in which this interpretation "engenders difficulties in dealing with contemporary pluralism"(Bohman and Rehg 32). Bohman and Rehg's concern is that Habermas invariably ties all political discourses back to a "singular set of idealizations." Most notable amongst these idealizations is the one which stipulates that, in the event that a discourse could be carried on indefinitely under conditions of full openness and inclusiveness that participants would converge on a discursively achieved agreement not for a host of diverse reasons but for the *exact same ones* (Habermas, Facts and Norms 166). For Habermas, the "concept of discourse seems to imply the possibility of full consensus toward something like the correct answer once ideal conditions are met" (Bohman and Rehg 33). If this is the case, then an ideal deliberative procedure could be able to, at least under counterfactual conditions of openness and inclusiveness, produce unanimous and full agreement between deliberating parties.

However, Bohman and Rehg argue that defining unanimity as the standard by which to judge rational agreement, especially when a conflict consists of a variety of different claims and where there appears to be no identifiable general interest on the horizon to help settle the dispute, rests on a questionable assumption about the *problems* that contemporary pluralist democracies grapple with, namely, that one "must assume that there are no intractable conflicts between different types of discourse" (43). What

Bohman and Rehg are referring to is the distinction Habermas draws between moral and legal norms:

Valid moral norms are “right” in the discourse-theoretic sense of just. Valid legal norms indeed harmonize with moral norms, but they are “legitimate” in the sense that they additionally express an authentic self-understanding of the legal community, the fair consideration of values and interests distributed in it, and the purposive rational choice of strategies and means in the pursuit of policies. (Facts and Norms 156)

As this passage suggests, legal norms involve many different kinds of considerations.

However, Bohman and Rehg are concerned with how some of these different considerations regularly come into conflict with each other, making unanimous agreement on legal norms unachievable. Similarly, McCarthy observes that there can be considerable disagreement even amongst the most well-intended parties to a moral discourse that are “not about...competing particular interests but about what ‘really’ is in the general interest” (McCarthy 196). If we grant both of these points, it is difficult to see how one could assume that all moral conflicts in society are resolvable. For, even if all deliberating parties argue with a reasonableness that mirror all of the traits that amount to a tacit acceptance of (U)—that is, they put forward good reasons to support their validity claim, consciously avoid making irrational or untenable claims, do not exercise any form of coercion over their interlocutors, and are receptive to the arguments of others and the modifying effects they may exert on their own position—there is still no guarantee that this will result in an agreement, let alone a consensus about the specific nature of the problem that is being discussed.

Linking the legitimacy of legal norms to the expectation of unanimous agreement makes a second further assumption about deliberating *citizens*, namely, “that citizens can always clearly separate the rational-discursive aspects of issues from those

aspects that require compromise.” (Bohman and Rehg, 44). This is a limitation that “may undermine the ability of discourse theory to deal with contemporary pluralism” (44). In his discourse ethics, Habermas holds that no speaker, “as long as he takes a *performative attitude*, confronts normative claims to validity seriously, and does not objectify norms as social facts, i.e., avoids reducing them to something that is simply found in the world,” can avoid the “claim of the principle of discourse ethics to transcend all local conventions” (“Discourse Ethics” 105). The problem posed by political and legal discourses, at least insofar connecting them in some way with discourse ethics, is that their concerns expand beyond generalizable interests (those that make “absolute” validity claims on speakers) to include the justification of norms with only a “relative” validity (relative in that their validity that can determined only against background “contexts” or “value horizons”) (Habermas, Facts and Norms 156). For Habermas, “policies and legal programs... do not raise moral questions *only*, but also involve empirical, pragmatic and ethical aspects, as well as issues concerned with the fair balance of interests open to compromise” (52). Notwithstanding the “complicated network of discourses and bargaining” that authorized political bodies engage in to mediate between these different aspects, Habermas nevertheless asserts that if a decision-making procedure for establishing political/legal programs is to enjoy a “presumption of fairness,” a presumption that confers legitimacy on such programs in the eyes of those affected by such programs, the procedures themselves “must be justified in moral discourses” (Facts and Norms 167). Bohman and Rehg remark that this postulation of an “internal connection”⁴ between normative and evaluative moments, between specifically moral and non-moral issues, implies that where “one cannot expect unanimity on substantive values

and questions of identity,” we still expect that participants could nevertheless adopt a “standpoint of impartiality and neutrality” by abstracting themselves from “particularistic sources of conflict,” thereby leaving “an opening within *moral* discourse for [the resolution of] disagreements resistant to argumentative resolution” (Bohman and Rehg, 45). In other words, despite the reservations expressed by some theorists that interpretations of justice are shaped by specific value horizons, Habermas still believes that a sufficient amount of impartiality and neutrality on the part of discursively acting agents will reveal common concerns between them (a generalizable interest) about justice. The problem, for Bohman and Rehg, with this belief is that:

In cases where sub-cultural ethical values and particular identities affect the very interpretation of justice and impartiality, then consensus would seemingly require different groups to resolve those differences behind their competing views of justice—a task one need not assume is possible in principle. Even aside from this problem, however, one may ask whether it is possible or reasonable to expect citizens to treat their moral identities as Habermas’s account of abstraction requires (45).

This point is significant, since it questions an idealization that lies at the heart of

Habermas’s moral theory: the requirement that moral agents take up an impartial and reflexive stance to their own value orientations.

Finally, making unanimous agreement/ consensus the standard of rational agreement amongst also assumes something about the *outcomes* of democratic deliberation. Idealized discourse has embedded in it the strongly epistemic notion that under ideal, constraint-free discursive conditions that the ‘force of the better reason’ will prevail, and that this entails convergence toward a correct, that is, a rationally justified, conclusion. Bohman and Rehg trace how this notion carries into Habermas’s procedural model in a way that implies “ideal convergence toward the correct answer...seems required even under non-ideal, actual epistemic conditions” (45). Accepting that ideal

convergence obtains even under imperfect conditions assumes that even incomplete deliberation, where discussion and reasoned argumentation are concluded if dictated by expediency or necessity—for were debate was to be prolonged indefinitely then nothing might ever get accomplished—, will necessarily lead toward something approaching a correct answer. Bohman and Rehg point out that given the finite nature of political conversation prior to political action we can expect that at least some of the parties involved in that conversation “may still have only an imperfect grasp of the relevant information and arguments” (45). Furthermore,

Incomplete deliberation, even if open and inclusive, does not necessarily increase the chances that the better argument will prevail. There is no guarantee under the less than ideal conditions of incomplete deliberation and information that there will be any view at all toward which the majority of citizens will converge, or if there is, that this view is the correct one.... Incompleteness by itself may make the deliberation underdetermining, even if good arguments are available. (45-46)

‘Incompleteness’ simply refers to how political decision-making proceeds under fixed deadlines, making it less likely that all reasonable positions on the issue under debate will have a chance to be heard, let alone thoroughly examined and considered. This is perhaps the most notable detail in pinpointing the dissimilarities between the idealized strictures of moral discourse and formal institutions for political deliberation, and suggests that an insolvent incompatibility lies somewhere between the two projects. Although Habermas undertakes to overcome this discordance by exposing the “dualistic structure” of political power—situating communicative power in the informal public sphere and administrative power in authoritative parliamentary bodies—the unsatisfactory association between the two is what motivates Scheuerman and Young’s concerns.

Simone Chambers expands on this notion of deliberative ‘incompleteness’ by claiming that discourse can only have a limited, but indispensable, function in assessing

not actual political decisions themselves but the quality of the deliberations that precede binding will-formation. Discourse, as a practice oriented toward reaching mutual understanding between speakers, can only be ascribed a limited role in assessing the rationality of binding political decisions because “short of an ideal communication community, [discourse] has no mechanism through which we can bring about closure” (Chambers, “Discourse and Democratic Practices” 235). This is because closure—defined as the necessity to eventually force, for reasons of efficiency, the discontinuation of debate so that a decision can be made—acts as an inescapable constraint on discourse, a practice that by definition is supposed to be constraint free. Chambers’ awareness that efficiency is an inexorable component of actual decision-making highlights the paradoxical relationship between discourses and democratic practice; that is,

The closer our conversations come to embodying the ideal [of unconstrained discourse], the more inefficient they are. The more general the norm under discussion, the more diffuse, fragmented and complicated will be the web of discourse, not to mention the longer the process is likely to take. With this in mind, it becomes difficult to even talk about a decision being taken in a discourse.... Indeed, the more parties to a discourse are constrained by the need to take a decision, the less motivated they will be to act discursively and the more motivated to act strategically. (250)

This analysis indicates that, as a tool for assessing the legitimacy of democratic will-formation, “discourse ethics should not be identified too closely with the day-to-day procedures of democratic decision and policy making” (250). Instead, because consensus-formation emerges only “partially” and is “fragmentary,”—as opposed to actual decisions that, in a democracy, are almost exclusively the direct result of majority vote—the critical standard of practical discourse should not concern itself principally with defining outcomes, but focus more on the institutions and procedures “where collective interpretations are constructed” (250).

2.3

In quite distinct ways, all the critiques so far discussed seem to point to how the stipulation of full consensus (consensus for the exact same reasons), which is explicitly brought out in Habermas's definition of rational agreement, has problematic side effects when brought to bear on political theory. As a way of securing the political conditions of publicity (impartiality and neutrality) necessary to bring about an argumentatively achieved consensus on validity (a consensus based solely on the force of the better argument), the political separation of informal public networks of communication from formal political bodies results in something of an awkward outcome, namely, a model of deliberative politics that *communicatively* empowers public bodies by depriving them of the *actual* power (although Habermas characterizes this as 'relief' from the 'burden' of decision-making) to steer political will-formation. And, as pointed out in the last section, the notion of a 'convergence toward consensus' relies for its plausibility on some impracticable expectations about political life. These expectations involve contentious claims about not only deliberating citizens (they can, or we should expect them to, distinguish 'abstract' moral issues from their deeper value commitments), but also about the kinds of problems those citizens deliberate and argue over (complicated and divisive issues are not intractable) and about the outcomes of those deliberations (on matters of justice a salient general interest emerges that can serve as a criteria of agreement by all and for the same reason).

While estimations of complexity and pluralism cast suspicion on the empirical veracity of Habermas's democratic project, these perspectives seem to say very little directly concerning the theory of normative validity that underwrites the version of

political legitimacy that their criticisms converge on. The problem with confining one's analysis of the discourse theory of democracy to questions of feasibility or tenability is they provide little in the way of any alternative to the communicative conception of the relation between basic rights and popular sovereignty. After all, Habermas utilizes discourse theory to expose an "internal relation" between private and public autonomy, and to show how they "mutually presuppose" each other, thus attempting to resolve a fundamental tension in the democratic theory literature between the two. If important insights like this one are to be retained, I argue, then some understanding of why Habermas's moral theory contains this ideal of consensus, what the problems are with this ideal, and whether or not his 'strong' version of consensus can be replaced with a less demanding one, is imperative. This question of 'replacement' becomes especially important in chapter three if the connection between normative justification and the motivating force of reason is to be retained.

As mentioned in some detail above, the idea that a rationally motivated agreement will result in an argumentatively achieved consensus on validity (i.e. that a valid moral norm reflects an agreement amongst all affected by it for the *exact same reason(s)*), is an expectation of practical discourses inferred from the principle of universalization (U). To repeat, as a rule of argumentation, (U) stipulates that a norm is valid iff:

All affected can *freely* accept the consequences and the side effects that the *general* observance of a controversial norm can be expected to have for the satisfaction of the interests of *each individual*. (Habermas, "Discourse Ethics" 93)

Thus, an expectation of consensus is not a rule of argumentation but part of the "logic of discourse;" that is, (U) defines a condition an agreed upon norm should meet if the rules of argumentation are observed. So, why does Habermas tie rational agreement to the

strong standard of absolute consensus, and why could rational agreement on a norm not be based on *different* reasons? The answer has a lot to do with Habermas's dispute with moral non-cognitivism, that is, the position that morality is no more than an expression of arbitrary desires and preferences that have no rational (objective) basis. Contrary to this position, discourse ethics is a *cognitivist* moral theory, not in the sense that it presupposes an "independent order of moral facts" existing in the world, but in that it ties "the acceptability of norms to the validity claims of regulative speech acts" (Baynes 110). On this definition, moral cognitivism takes on an intersubjective interpretation whereby "norms no less than assertions can be contested and the validity of both depends upon their capacity to be redeemed through discursive argumentation" (110). Whereas the validity of scientific claims can be adjudicated discursively by reference to the manifold of empirically supportable facts independent of social relationships with others, "normative claims to validity *mediate a mutual dependence* of language and the social world that does not exist for the relation of language to the objective world" (Habermas, "Discourse Ethics" 61). As such, valid norms are not fixed once and for all but "are dependent upon the continual reestablishment of legitimately ordered interpersonal relationships" (61). Since norms exist in a "universe of norms," and do not correspond to a fixed and independent state of affairs the way empirical statements do, normative justification can only occur within actual exchanges of reason—exchanges whose rules of engagement are determined by inescapable presuppositions of communicative rationality—wherever normative consensus breaks down.

Notwithstanding this difference between empirical truth claims and normative rightness claims, the validity of both kinds of claims depends on the existence of rules

that determine their validity. In terms of both scientific discourses and normative discourses, “the sense of a claim to truth... is revealed by examining the practices connected with making and defending truth claims: one understands the meaning of “truth,” as we use it in everyday parlance, when one can describe the rules for using the terms” (Rehg, Insight and Solidarity 26). For making empirical claims to truth, these rules pertain to methods of induction, deduction, falsification, and the gathering and giving of evidence. Likewise, justifying the validity of rightness claims involves creating rules for practical discourses based on reconstructing “the intuitive “know how” of competent speakers” (26). These rules outline the “validity conditions” under which someone can convince someone else “that he is entitled in the given circumstances to claim validity for his utterance—in short ...*what makes it acceptable*” (Habermas, “Actions, Speech Acts” 232). Meeting the conditions of a practical discourse in argumentative praxis, akin to following rules of empirical discourses, has the effect of bringing about a consensus amongst discursively-acting parties on normative claims to validity. It is this connection between the discursive redeemability of truth claims and rightness claims, ‘redeemability’ referring to the potential to reach consensus on both kinds of claims, that gives the “ought” of the latter kind of claim its normatively binding (motivating) force.

Seyla Benhabib seeks to soften Habermas’s strong reading of discursive consensus, in which rational agreement on norms is tied to an outcome that convinces all for *identical reasons*, by questioning whether or not discourse ethics can jettison the controversial premise of (U) yet retain a “cognitively articulable kernel.” The rationale behind this approach is Benhabib’s assertion that for communicative ethics to become a

“concrete” moral theory, rather than one that is “too indeterminate” or “too counterfactual” for making actual moral judgments, it must not restrict itself to the narrow range of questions afforded by matters of justice but also be capable of thematizing and addressing questions of the good life. Benhabib begins by suggesting that (U) is a “redundant” principle, and that in stipulating consensus as a necessary condition of rationally motivated agreement (U) “adds little but consequentialist confusion to the basic premise of discourse ethics” (“Afterword” 344). This appraisal springs from her own competing and alternative interpretation of the unavoidable presuppositions of argumentative speech. Whereas Habermas argues that as a moral principle (U) “is implied by the presuppositions of argumentation in general,” Benhabib derives two different principles from those same presuppositions: universal moral respect and egalitarian reciprocity.⁵ The crucial difference between these two principles and (U) is that they do not explicitly express a specification which “every valid norm has to fulfill,” a specification that indirectly restricts what the *result* of a moral discourse should be. Rather, what these two principles do is reflect, based on the rules of argumentation, the substantive ethical preconditions already built into the *process* of moral discourse while remaining silent about ends, content, or outcomes. Universal moral respect and egalitarian reciprocity do not abandon the rules of argumentation that Habermas sees (U) as an extension of, but merely recasts these rules as unavoidable presuppositions of symmetry and reciprocal recognition built into the language of ethics. The advantage of reading communicative ethics in this way is that it both retains the idea of norms as more than statements of “preference” or “taste” and abandons the problematic notion of an “appropriate moral domain” delimited by the requirement of strong consensus.

Joseph Heath also questions whether or not the stipulation of consensus is a necessary element of discourse ethics. Heath views the theme of 'convergence toward a consensus' in Habermas's moral theory as an "unnecessary burden of proof" that discourse ethics shoulders in defending the rationality of moral argumentation from the claims of moral noncognitivism, further arguing that cognitivism and convergence are notions quite distinct from each other. Heath's problem with the consensus aspect of Habermas's theory, as we will see, is not so much that it is impractical or unworkable in practice, but that the equation of moral justification with an argumentatively achieved consensus involves some questionable and, in the end, unnecessary lines of reasoning. Furthermore, when taken together, these claims reinforce the even larger and more contentious theoretical postulate in Habermas's work which occupies a large part of Heath's examination of discourse ethics, namely, Habermas's argument that a rationally motivated agreement cannot involve elements of bargaining and compromise. That issue will be taken up in chapter three.

Heath's interrogation of these unsound claims begins by way of an analysis of Habermas's moral-ethical dualism. According to Heath, discourse ethics' status as a moral cognitivist theory relies on a connection between conditions of moral argumentation and rules of inference; that is, the rules or conditions under which a moral validity claim could be discursively redeemed (justified) establish the counterfactual assumption that "*in principle...* rationally motivated agreement must always be possible" (220). The result of establishing this connection is that it "ties the cognitive status of moral judgments to an a priori expectation of convergence in moral argumentation" (220). Such an expectation subsequently commits the discourse ethics program to the

conclusion that given enough time for the intersubjective give and take of reasons, and the elimination of any external and internal constraints on discursively acting parties, that there could not be an intractable disagreement between participants involved in a moral discourse.

While the expectation that, in principle under ideal conditions, a convergence toward full moral agreement is possible is not problematic on its own, Heath argues that the resulting direction Habermas takes his inquiry, specifically his explanations of moral disagreement, is less convincing. Habermas confines his explanation of moral disagreement to two sources, either that the claim being debated is not a strictly moral norm but involves an ethical claim that is intertwined with a particularistic value orientation not shared by all subject to the norm under consideration, or, the conditions of openness and inclusiveness under which discursive justification takes place are not being fulfilled. According to Heath, (U) is not part of the 'logic of discourse' that follows from the rules of argumentation, but instead acts as an "ancillary device" designed to ground a moral-ethical distinction which makes those two accounts of moral disagreement more credible while also supporting the idea that full moral agreement is possible. More importantly, in servicing the analogy between moral argumentation and empirical inquiry, (U) serves as the rule of inference for warranting the rational acceptability of moral norms that functions in a fashion similar to the way induction does for the construction of scientific theories. Just as by way of induction a general theory can be derived from the inexhaustible manifold of observable facts, Habermas asserts that the universalization principle, a principle which outlines the acceptability conditions for the outcomes of any practical discourse, "bridges" the gap between a norm that reflects a

general interest amongst all those affected by the norm and the plurality of different interests and value orientations.

Human interests, however, differ from empirical facts in several ways that make it difficult to believe, for three reasons Heath offers, that normative conflict can, in principle, be resolved by appealing to a universal “common interest” (an interest presupposed to be held by all potentially affected by the norm under consideration). First, there is no convincing reason to think that norms thought to lie in the common interest may not come into conflict with each other. In fact, it is often when two generally accepted norms come into conflict that moral argumentation is precipitated at all. Second, isolating a common interest from amongst a constellation of diverse interests held by the participants to a discourse is not an easy task. Hence, Heath is perplexed by Habermas’s endorsement of, in spite of his commitment to dialogical justification, what seems to be an intuitive guarantee of convergence toward a moral consensus rather than understanding moral agreement as “only ongoingly accomplished, in ever-changing circumstances, and for all practical purposes” (Hoy and McCarthy 244). Third, even in the unlikely event that there exists a unified common moral interest amongst people of different socio-cultural backgrounds, and the even unlikelier event that such an interest could be identified, such a common interest might be too thin to provide action guidance in thicker sociocontextual-specific conflicts. For instance, a common interest in the inviolability of all human lives has done little to settle the deep-rooted conflict between pro-life (anti-choice) and pro-choice (anti-life) viewpoints in the debate over access to abortion.

That a universally-held, generalizable interest should be the source of normative justification in discourse theory is interpreted by Heath as symptomatic of Habermas's attempt to "banish *values* from the domain of moral argumentation, because he regards differences over questions of value as the primary obstacle to agreement" (Heath 234). In effect, with this statement Heath is simply turning Habermas's own commentary on Rawls—that the original position preemptively and unwarrantedly "neutralizes the multiplicity of particular interpretive perspectives from the outset,"—back toward discourse theory (Habermas, "Reconciliation Through" 57). For Heath, this 'general interest/ particularistic value' distinction hinges on Habermas's stance that participants to a discourse can abstract themselves from their value orientations in a way that they can uncover in dialogue with others a mutual interest, a capacity for abstraction which Heath rejects based on the basic hermeneutic premise⁶ that an agent's values are exclusively what their own interests and those of others are interpreted through. This criticism resonates not only with Bohman's and Rehg's earlier remarks about how ethical identities shape an agent's interpretation of justice, but also with Benhabib's own conclusion that "even from within the constraints of a discourse theory, [the] hard distinction between judgments of justice and those of the good life cannot be maintained" (Benhabib, The Claims of Culture 358).

That values cannot be usefully excised from more general interests in a more convincing way has further repercussions for the distinction Habermas draws between moral and ethical discourses. For, if participants to a discourse cannot engage in the kind of abstract reasoning that allows them to "step back from all contingently existing normative contexts," it makes little sense to further postulate discrete ethical

commitments where participants “cannot work themselves free from the form of life in which they de facto find themselves in” (Habermas, Facts and Norms 163). Holding fast to this moral/ethical distinction, particularly when liberal democratic societies are witness to an ever increasing amount of diversity and value orientations within the populace, has the tendency to “shrink the sphere of questions that can be answered from the moral point of view” (Heath 238). Thus, a liberal society’s gain in value pluralism, a trend against which there is no convincing argument, threatens to “sever the connection between Habermas’s discourse ethics and his general theory of communicative action” (238), subsequently making discourse ethics an increasingly less productive theory with which to ground the legitimacy of democratic decision-making. This is especially the case if Habermas ties political discourses back to a singular set of idealizations in the way that Rehg and Bohman thinks he does.

Heath does agree with Habermas, though, that there is an important action-theoretic distinction between norms and values that needs to be maintained, namely, “that norms should not directly institutionalize values... [since] norms are public in a way that preferences are not” (241). For, it is this ‘public’ character that makes norms binding and gives them “the absolute sense of an unconditional and universal obligation;” by contrast, values “compete for priority... to the extent that they find intersubjective recognition within a culture or form of life” (Habermas, Facts and Norms 255). However, establishing social norms does not require exclusion at the outset of particular evaluative preferences for consideration as a valid social norm, since it is only in the tabling of our private values in a public forum that we can jointly determine whether these values are shared with others, for what reasons, and to what extent. Conversely, we cannot force our

values on those that do not share them either, for that undermines the idea that a valid norm is one that is followed on the basis of good reasons. So, in cases where some shared values already exist, they can be appealed to in establishing a general norm. In cases where there is no readily identifiable shared value upon which to base a norm, it is less clear what we appeal to in grounding a moral argument or settling a dispute over a norm

If disputation over a norm cannot be brought to resolution by an appeal to a shared value amongst the parties concerned, then how can one use the idea that rational agreement does not necessarily involve all parties agreeing for the same reason to establish an alternative means of for resolving normative conflict? When there is no “common metric of value,” Heath claims, the best way to potentially settle a conflict of interests in the establishment of a social norm is to bargain so as to bring about some sort of a “reasonable compromise” amongst all affected by a norm. This suggestion is in my view a powerfully sensible one because it seems well suited to dealing with the incommensurability of values and different life projects that mark liberal democratic societies. Furthermore, as chapter three will attempt to elucidate, a broader understanding of compromise formation has a great deal of potential to ameliorate elements of Habermas’s democratic theory in a way that both follows his theoretical perspectives (although in a somewhat modified form) and makes it more able to respond to the criticisms of his political theory brought up in this chapter. In drawing out the implications of this broader understanding, however, there remains a difficulty for dealing with what is the rather limited function Habermas assigns to compromise formation and bargaining procedures for generating stable democratic societies. This view, according to Heath, is largely informed by a mistaken assumption Habermas makes

about compromise formation: that compromises always involve the use of purposive-strategic, rather than communicative, rationality, and that as such compromises do not involve the achievement of a mutual understanding but are always a balance of conflicting interests.

Chapter Three: Discourse, Compromise and the Public Sphere

For Habermas, bargaining arrangements in democratic societies are reserved for those times when, as is “often the case,” deliberating parties can neither “submit their contested interests and value orientations to a universalization test within the framework set by the system of rights as it has been constitutionally interpreted and elaborated,” nor can they “push beyond contested interests and values and engage...in a process of self-understanding by which they become reflectively aware of deeper consonances in a common form of life” (Facts and Norms 155). And since legal programs and statutes usually involve a combination of moral and non-moral (non-universalizable) interests, it can be reasonably expected that for complex societies many of its “proposed regulations [will] touch on the diverse interests in respectively different ways without any generalizable interest or clear priority of some one value being able to vindicate itself” (165). Bargaining arrangements and achieving compromise seem to be a political necessity for resolving conflicts when the prospect of reaching unanimous agreement is remote. But if unanimity is the discursive standard of a rational agreement, what gives political will-formation the rational character Habermas thinks it can have if such a strong agreement is unworkable in practice? According to Habermas, the legitimacy of compromise formation is tied to the *fairness* of the negotiating procedures themselves, and to have the presumption of fairness these procedures must maintain an “indirect” association with the discourse principle. “More specifically, the negotiation of compromises should follow procedures that provide all the interested parties with an equal opportunity for pressure, that is, an equal opportunity to influence one another

during actual bargaining, so that the affected interests can come into play and have equal chances of prevailing” (165).

While the idea that a legitimate compromise is one that is fairly negotiated is a reasonably intuitive stipulation, Habermas thinks that bargaining arrangements “are tailored for situations in which social power relations cannot be neutralized in the way that rational discourses presuppose,” and that these arrangements must “involve success-oriented parties who are willing to cooperate” (166). There are three controvertible claims about compromise formation embedded in this description; first, this description has a very restricted definition of when bargaining arrangements are required to settle conflicts; two, the involvement of “success-oriented parties” in this description intimates that compromises are strategic rather than communicative interactions; and third, in this description compromises do not involve parties striving for “mutual understanding” but simply provide an equal opportunity for each party’s interests to “prevail,” which suggests that, from the perspective of a practical discourse (where agents seek to reach a mutual understanding), negotiated compromises are less rational than an argumentatively achieved consensus on validity. Moreover, as pointed out in the introduction, when it comes to resolving conflicts of interests, Habermas subordinates bargaining and compromise as a reasonable means for dealing with conflicts of interests to the achievement of consensus in practical discourse.

As was demonstrated at the end of chapter two, the notion of a unanimous consensus, even on basic questions of justice, is for some critics a questionable ideal of rational argumentation at best and seems unnecessary at the dialogical level of moral theory. Nevertheless the notion of consensus also seems to indirectly inform Habermas’s

approach to compromise formation. This is regrettable since it prevents him from working out a discourse-theoretic approach to democratic compromises and its role in his larger picture of political legitimacy. In what follows, I will draw on two analyses, one based on rational choice theory and the other a more nuanced account of public political deliberation under conditions of value pluralism, to repudiate Habermas's view of compromise as a strategic form of action coordination, thus bringing out the important role compromise formation plays in establishing stable, just and inclusive democratic societies. As we shall see, however, the idea that normative validity involves full consensual agreement must still be retained, albeit in a modified form. After establishing the possibility of nonstrategic compromises (settlements of conflict that involve the exchange of 'non-shared' communicative reasons), I will conclude by making some suggestions, in light of James Bohman's notion of a 'moral compromise,' as to how compromise formation can be integrated into the discourse theory of democracy in a way that allows it to reply to its critics without imperiling its foundations in moral cognitivism.

3.1

Game theory refers to models used to understand and illustrate how rational choice—defined as an individual's orientation to make choices which optimize the utility that can be expected from a particular course of action—works when individual decisions are linked to interactions with other similarly rational individuals. In cases when an outcome of a decision is solely determined by one individual's actions, choosing rationally means choosing the action that maximizes utility. Games are different, though,

because they involve situations “in which the action of one person perceptibly affects the welfare of another and vice versa” (Hargreaves Heap and Hollis, 94). In other words, making a rational decision in a game that involves two agents requires that before an individual acts, they must first consider what the other individual is going to do, for the action of each has direct bearing on the outcome that obtains. One game in particular, usually referred to as the ‘Prisoner’s Dilemma,’ is effective in showing how instrumentally rational individuals—individuals with particular preferences who will only choose actions that they think will best satisfy those preferences, also known as a *utility maximizer*—can act in ways that generate suboptimal and collectively self-defeating outcomes in situations where the outcome is tied to the choices of another equally rational individual. The prisoner’s dilemma illustrates how when an outcome depends on the actions of two perfectly rational individuals, the result can be a decidedly irrational outcome, one where both individuals are worse off than had they chosen to act as something other than a utility maximizer. In these cases, two individuals acting instrumentally bring about an inefficient and therefore irrational collective outcome, a paradox of rational choice theory referred to as a *collective action problem*. One solution to this problem, the Nash bargaining solution, is where Heath feels a more convincing explanation of the rational acceptability of moral norms lies and that circumvents the problems associated with the existence of a common general interest.

The typical version of the Prisoner’s Dilemma features two players, both of whom are criminals, arrested for the same crime upon which there is not enough evidence to convict either on the most serious of the charges but enough to convict them of a less serious one. They are both brought to a police station and placed in separate interrogation

rooms. The two criminals neither conferred in advance about what to do should they find themselves in this situation, nor will they be given a chance now. The logic of the investigator tells her that these criminals are rational individuals and, when given the opportunity, will take the chance to maximize their individual expected utility. They are given two choices: to confess to the more serious crime or to not confess to it. If one confesses and the other does not, the confessor will serve one month in prison (the best possible individual outcome) and the unforthcoming one will receive a jail term of 10 years. If both confess, they will each receive 7 year terms. If confesses neither, they both receive three year terms. This leaves four possible outcomes (in order from maximum possible utility to minimum possible utility): one month, three years, seven years and ten years. The dilemma is that neither knows what choice the other intends to make, but as two individuals looking to maximize their utility both will act strategically, choosing to confess because only by making that choice does either have a chance at achieving their individual best outcome (one month imprisonment) and avoiding the worst individual outcome (10 year imprisonment). While by confessing they both avoid the worst outcome, they would clearly have both been better off choosing not to confess and receiving three year sentences (the second best outcome), rather than the seven years (second worst outcome) they received by both confessing, which is still an irrational outcome given that they both have the same interest in maximizing their utility.

It might seem that a simple way to overcome this collective irrationality would be to take the agents out of isolation so they can *talk* to one another and engage in rational negotiations. However, one basic element of bargaining theory is that there needs to be some incentive to cooperate, and when agents are taken out of their symmetrical

positions and begin talking, each agent quickly starts to learn things about the other that reduces the incentive (from an instrumental point of view) to cooperate with each other. One or both may discover disparities between each other in terms of knowledge, power, position, risk aversion and time that equip the agents with *bargaining power*, an advantage that gives agents greater ability to steer the eventual outcome, which can then be brought into the negotiations. And since we continue to suppose that both agents have the sole interest of maximizing their utility, this power reduces the incentive to act cooperatively and increases the incentive to act strategically and non-cooperatively. However, as the Prisoner's Dilemma shows, this strategic orientation can lead to an irrational outcome, the most irrational outcome being that no agreement is reached at all. And in bargaining theory, reaching some kind of agreement, where at least some of the available utility is made use of, is rationally preferable to not reaching any agreement at all. This is not to say that agreement through bargaining is always possible 'in principle,' but only to "suggest that disagreement can be expected to persist under conditions of very extreme moral or discursive intransigence" (Heath 244).

A bargaining situation is defined "as one in which individuals 'have the opportunity to collaborate for mutual benefit in more than one way'" (Heath 244). In light of the difficulties and impracticalities associated with discourse ethics' postulation of a common interest or common sense of well-being, rational bargainers in Nash's solution "care only about the utility they will get from agreed settlements... [And] considerations such as risk aversion, envy, sympathy, concern for justice and so on are all supposed to be included within the function that converts utilities into payoffs" (Hargreaves Heap and Varoufakis 135). Nash's generic bargaining problem features two

players seeking to reach an agreement on how to split between them the dollar they have together just found on the street. Terms of the agreement are discussed and debated within a fixed amount of time,⁷ and then each submits a bid for the portion they want of the dollar.

Constructing a bargaining situation first requires establishing the *disagreement point* of the conflict in question as well as a *bargaining space*. The bargaining space simply refers to “the set of possible cooperative outcomes in terms of their payoffs” (Heath 245). Within this space there are many different possible agreements, each of which represents some combination of utility that would make both agents better off than had they failed to reach agreement at all. The disagreement point, also known as the *noncooperative outcome*, exists at one corner of this bargaining space and refers to the outcome that occurs when agents fail to reach any agreement at all. At the disagreement point the potential utility of agreement (the entire dollar) is squandered. In the generic bargaining problem, the noncooperative outcome is the result of both agents acting strategically rather than cooperatively and is the situation rational bargainers are most interested in avoiding.

Nash’s solution to the bargaining problem is *axiomatic*, meaning that the solution is based not on an “analysis of actual bargaining but, instead, on some principles which he suggested any agreement between rational individuals should satisfy” (Hargreaves Heap and Varoufakis 130). These axioms work as “intuitive constraints” that any bargaining solution should fulfill. The first of these axioms is *efficiency*, the principle that any bargaining solution “should not involve any waste, or inefficiency, as rational bargainers would never agree to throw part of the [dollar] in the dustbin, rather than

distribute it amongst themselves” (147) Nash’s second axiom refers to the *symmetry* that any bargaining solution must have. Because we are assuming that the two bargainers in question have the same desires, motives and information, any differences between them not directly pertaining to their rationality as bargainers should have no effect on the outcome. A symmetrical outcome means that the utility distribution of the solution should not be such that one player would want to change places with the other and thereby change the agreement.

In and of themselves these axioms are not very significant because they “follow fairly directly from simple “role-taking” operations and the two constraints can be derived from the structure of practical discourse” (Heath 246). These axioms only indicate rational solutions where there is no difference between the players. However, as Heath rightly notes, “people are usually in asymmetric positions with respect to particular interactions, and so will have different opportunities and preferences” (246). The third element of axiomatic bargaining solution is known as the *independence of irrelevant alternatives* (IIA), and, Heath explains, shows why such differences in relative bargaining position should have no bearing on the initial outcome. The IIA “entails a consistency requirement across different bargaining solutions and requires that a solution to a bargain should not depend on the set of excluded outcomes” (Hargreaves Heap and Varoufakis 147).⁸ For Heath,

The significance of this axiom is that, when satisfied, it implies that for any given distribution problem there will be a perfectly symmetric problem that is given from the moral point of view. Since the two problems are equivalent, the solution that we would pick for the symmetric problem should also be the solution to the asymmetric problem. Thus we can “project” the solution from unproblematic cases onto more problematic ones. (Heath 246)

This third axiom is more contentious than either the efficiency or symmetry requirement, because it seems irrational that, were the initial bargaining situation to change in a way in which an agent could act differently (perhaps because they have acquired information that elevates their relevant bargaining position) to increase their utility from the interaction, they would not act on this. Conversely, Heath thinks this objection is salient only if the players involved in the game are acting strategically, which, according to the parameters set out in the generic bargaining problems, would result in an outcome that would violate one of Nash's other two axioms. It is by no means clear, however, that Heath's explanation overcomes this objection, since what constitutes an irrelevant alternative (as opposed to a relevant one) for an agent is largely contingent on what counts for each agent as 'relevant.'

Nash's bargaining solution is only a "skeleton" analysis of rational bargaining insofar as it merely outlines how well-informed bargainers engage each other in trying to reach a negotiated agreement. The solution is simply an attempt to identify an "already existing reason" in bargaining situations. Of course, in real negotiations, not all negotiators are identical and some participants have resources which others lack that can be used to strategic advantage in their interactions. Such disparities will sometimes lead to unfairly negotiated agreements or, sometimes, to no agreement at all. Nevertheless, Heath uses Nash's solution to draw out what sees as two fundamental misunderstandings in Habermas's work concerning bargaining processes and compromise formation.

The first misunderstanding is that bargaining is a necessarily strategic form of action, a conception Heath thinks is motivated at its foundation by the "problematic

assumption...that when [agents] are unable to achieve consensus on a common interest or value, they have no choice but to drop the performative stance and adopt a strategic orientation” (Heath 248). The problem with this picture is that, as the axiomatic bargaining solution indicates, a negotiated compromise does not require a consensus amongst agents on values or interests. The only consensus (although in a weaker sense than how Habermas uses the term) required, though still brought about through discursive means, regards whether or not the compromise fits with the axioms. Only a player’s “acceptance of the relevant set of axioms gives them good reason to accept the solution that uniquely satisfies these constraints” (248). Furthermore, it is by no means clear that agents accept outcomes that conform to the axioms in the bargaining solution for instrumental reasons, since there are always outcomes outside the bargaining solution in which a given agent’s expected utility will be higher. And, Heath thinks that “since [agents] will each have an instrumental incentive to pursue this outcome, insofar as they accept the bargaining solution, it cannot be for instrumental reasons” (249).

The second misunderstanding about bargaining comes from what Heath labels Habermas’s “deep confusion” about compromise formation. For Habermas, linguistically mediated interactions can be categorized into two “elementary types of action,” communicative and strategic action, “neither of which may be reduced to the other” (“Actions, Speech Acts” 220).

In the [case of communicative action], the consensus achieving force of linguistic processes of *reaching understanding*—that is, the binding and bonding energies of language itself—becomes effective for the coordination of actions. In the [case of strategic action], by contrast, the coordinating effect remains dependent on the *influence*—functioning via nonlinguistic activities—exerted by the actors on the situation and on each other. Speech acts cannot be carried out with the simultaneous intentions

of reaching an agreement with an addressee with regard to something and exercising a causal influence on him. (220-221)

In the same way as he draws this distinction between elementary linguistic action types, Habermas applies this communicative/strategic dualism to the democratic decision-making processes of complex societies when he distinguishes between discursively reached public (informal) *opinions* and the binding (formal) *will* of those bodies charged with taking actual decisions. “In the former case, appeal is made to the consideration of norms and values; in the latter, that of interest positions” (Habermas, Facts and Norms 140). Heath breaks this distinction between communicatively oriented opinion-formation and instrumentally orientated will-formation down to the “different justificatory reasons” one considers in reaching agreement, a distinction that Heath then offers some persuasive reasons to question. For instance, Heath asserts that there is “no reason to think that simply because different agents find different arguments compelling, there is some strategic dimension to their actions” (251). Under the conditions of the axiomatic bargaining solution, different arguments for the solution to a conflict are acceptable so long as the solution meets “reasonable constraints” for “a cooperative solution,” constraints that agents determine by engaging in a practice “of ‘ideal role taking,’ to confirm that any other proposal would generate reasonable objections from some quarters” (251). Thus, the “only reason for finally accepting a compromise is precisely that it is a compromise,” and Heath considers that “it is the rational insight that the compromise is acceptable that provides the agent with a reason to accept it” (251-52).

3.2

Rational bargaining, as Heath's analysis shows, is more than just a balance of conflicting interests. It presupposes that participants implicitly accept discursive constraints of reasonableness (symmetry, efficiency) on proposed solutions to conflicts that are similar to the structures of a practical discourse. 'Fair compromises' involve not the exchange of strategic reasons that can then be "discursively regulated" or filtered through 'fair' procedures that are justified in specifically moral discourses, but the exchange of non-identical communicative reasons. This exchange still requires agents to engage in some form of 'ideal role-taking' when reflecting on any proposed solution, a process that Heath does not really describe in any detail. Furthermore, agents do not accept a compromise for different reasons but for one reason, namely, that any alternative to a fair solution is an outcome that will prove unacceptable to at least one party involved. But, regardless of how rational the axiomatic bargaining solution looks from the perspective of game theory, there still remains lingering doubts about just how 'nonstrategic' this kind of bargaining is. These doubts, which I will bring attention to shortly, become all the more acute when seen through Heath's elaboration of the 'pragmatic' explanation of convergence on social norms.

Like Habermas, Heath does not think it possible to explain a stable social order by recourse to purposive-instrumental rationality alone, since "instrumentally rational agents cannot credibly commit themselves to future actions, even when the adoption of such commitments would benefit all" (88). This problem is manifest in the earlier discussion of the Prisoner's Dilemma, a game whose contours illustrate how agents acting instrumentally rational (in a way oriented to optimizing individual utility), as opposed to cooperatively, can bring about a suboptimal or inefficient outcome.

Furthermore, as Heath points out, the instrumental conception of rationality is at a loss to explain “the fact that agents *do* routinely make and honour promises, and that these commitments provide the foundation for the development of all large scale cooperative activities, including agriculture, industry, construction, and navigation” (89). However, against Habermas, Heath argues that the “accountability of social action cannot be a product of the accountability of speech acts” (127), that is, he “rejects the idea, central to Habermas’s view, that social norms acquire their binding force from an associated rightness claim” (170).

Consider Heath’s assertion that people converge on social norms because of the “pragmatic benefits” associated with doing so, namely, the avoidance of indeterminate and inefficient outcomes. Remember that the discussion of nonstrategic bargaining began with the assertion that we should not concern ourselves too much with how rational people agree (what Habermas is concerned with when he tries to show that agreement is always possible in principle), but be more attentive to those instances where rational people *disagree* over norms. For Heath believes a state of affairs in which a society does not for the most part converge on common norms for their shared life is only possible when rational agents are “extremely disagreeable.” The expectance of convergence on moral norms is for Heath not, as it is for Habermas, built into the structure of practical discourse itself, but has a more *pragmatic* basis of which I will only highlight the features salient to this discussion.

In the generic bargaining problem, the disagreement point represents a conflict where the potential utility that could come from reaching any agreement whatsoever is thrown away entirely. Similarly, strategic actions predictably decrease the likelihood that

the available utility will be divided symmetrically and efficiently distributed, which means that inefficient outcomes are less likely to be enduring and stable than efficient ones because there will continue to remain 'fairer' alternatives that at least one agent (or many agents) would prefer. Applying this same line of reasoning to explain high levels of convergence on social norms, Heath explains:

There will be strong pragmatic reasons for maintaining agreement on a set of shared norms, since the indeterminacy of strategic reasoning ensures that the stability and predictability of social interaction will break down in their absence. Persistent disagreement over normative questions leaves agents unable to coordinate their interactions through institutional mechanisms. This leaves them only strategic action to fall back upon, an option that is often unattractive from both a moral and instrumental point of view. Not only is strategic interaction usually indeterminate...but it also harbors significant potential to produce suboptimal interactions. When this is likely to occur, agents are in a position to recognize that they would all be better off if they could come to some agreement. (272)

The point of this passage is that an explanation of convergence on social norms does not have to rely on the presumption of a common interest implicit in the structure of a practical discourse, but can be explained solely by the need for a 'stable' and 'predictable' social environment in which people are able to act in accordance with their values, regardless of what they may be and to the greatest extent possible.

Although the pragmatist account of convergence provides a plausible description for why there is such wide agreement on social norms, Heath's explanation that convergence on moral norms is tied to the "pragmatic benefits" associated with doing so has also come under some scrutiny. Arash Abizadeh contends that Heath's use of the axiomatic bargaining solution to explain normative action, while rightly broadening the range of agreements (to include compromises) that are morally justifiable, is nevertheless "tacitly parasitic" on the universalization principle. In contrast to some of the authors discussed in chapter two, and though he considers the problematization of Habermas's

moral-ethical dualism understandable, Abizadeh thinks that (U) nevertheless plays “an indispensable critical function,” albeit a different function than Habermas ascribes to it, in the discursive theory of moral justification. The critical function being referred to here is (U)’s capacity for “distinguishing the regulative ideal governing moral validity from any particular rationally motivated agreement on an articulated norm” (Abizadeh 194).

Abizadeh rejects Heath’s argument that (U), because it specifically defines moral norms as somehow embodying a set of general interests, “cannot be defended without reintroducing a certain foundationalist line of reasoning” (Heath 231).

Conversely, Abizadeh asserts:

It is quite true that (U) purports to justify norms in terms of interests. And so one might say that norms are “founded” in interests. But this is not foundationalist in the relevant sense of appealing to *premises whose truth is non-inferentially justified*, which is precisely how Heath himself defines foundationalism... what is being identified here is *not* any particular substantive factual premise whose truth must be taken as self-evident in some way: (U) says that, to be valid, norms must be justified in terms of interests; it does not say that norms must be justified by reference to [any] self-evidently true fact. (201)

Furthermore, Abizadeh thinks that (U) is an essential component of discourse ethics since it acts as a *constraint* on moral justification; specifically, (U) acts as a *welfare* constraint that “implies that moral validity is conceptually linked to a theory of value one of whose components is human welfare” (202). Significantly, this welfarist constraint does not state that human welfare is the *only* value appealed to in justifying a norm (meaning that other values could also be appealed to), nor does it specify exactly what human welfare consists of, as these are both issues that need to, themselves, be taken up in a practical discourse.

Demonstrating this conceptual link, however, still requires some explanation of how:

The pragmatic presuppositions of discourse commit any moral agent who uses language communicatively not just to the pursuit of normative agreement but also to a theory of value that (at least) recognizes human welfare as a genuine value (which thus makes a claim upon us). (203)

It is not Abizadeh's intention to supply that explanation upfront, but to instead provide an alternative reading of (U) that implies something like a welfarist constraint implicit in the "meaning of normative validity." Unlike Benhabib and Heath, both of whom think of (U) as little more than an unnecessary step Habermas takes in refuting moral noncognitivism, Abizadeh sees (U) as a "critical principle for marking the defeasibility of any rationally motivated consensus in a practical discourse as specified by (D)" (205). Up until this point in the discussion, (U) has primarily been interpreted as a misguided effort to "tie the cognitive status of moral judgments to an a priori expectation of convergence in moral argumentation" (Heath 120). However, Abizadeh views the suggestion that moral norms must somehow reflect a generalizable interest as capturing not only the meaning of normative validity, but also as an apprehension of the inherent "fallibilism of moral justification." On this reading, a common interest is not a "telos" to be realized through practical discourse, but just articulates what a consensus might look like "in the first place," and thus "provides...a way in which *any* articulated consensus can be called into question by exposing its necessarily contingent (i.e. defeasible) nature" (Abizadeh 207).

This is not exactly a novel interpretation of (U), but an important one nonetheless shared by other philosophers. For example, Rehg asserts that one way to look at (U) is to see it as "weakly conclusive": that is, (U) just stipulates that the dialogical justification of norms should always be viewed as fallible, and open to revision in light of new information, arguments, and points of view not previously considered (or now

considered in a different way). This stipulation, though, need not diminish any considered conviction that practical discourses, and a serious commitment to argumentation, can bring about an argumentatively achieved consensus on a moral norm, since in accepting this reading of (U):

We need not exclude the possibility that in *some cases* we might consider our grasp of an issue, and our arguments, as secure enough to make it highly unlikely that any future additional information could overturn our current consensus. (Rehg, "Discourse Ethics" 89)

In other words, when agents commit to the dialogical justification of moral norms they are also committing to the strong standard for justification expressed in (U), even though that standard most likely will never obtain. The significance of Regh's analysis here is his conclusion that while all moral justification has an ineradicable element of contingency, this need not undermine a commitment to consensual dialogical justification.

Surprisingly, a slightly different gloss on this 'defeasibility' reading of consensus can also be found in Benhabib, who believes that:

Discourse theory develops a normative and critical criterion by which to judge institutional arrangements, insofar as these current arrangements suppress a 'generalizable interest.' One can use this critical criterion as a critical yardstick by which to uncover the underrepresentation, the exclusion and silencing of *certain kinds* of interests. In other words, it is not so much the identification of the "general interest" which is at stake, as the uncovering of those partial interests which represent themselves as if they were general. ("Afterword" 352-53)

As discussed in chapter two, Benhabib drops Habermas's definition of (U) in her version of communicative ethics and replaces it with the principles of 'universal moral respect' and 'egalitarian reciprocity,' both of which she believes follow from the pragmatic presuppositions of argumentative speech. Even though Abizadeh's own opinion, that (U) plays an "indispensable critical function" in Habermas's discourse ethics, is stimulated by Benhabib's formulation of a critical potential in communicative ethics that explicitly

rejects (U), their positions are not automatically incompatible. Abizadeh simply thinks that (U) is not intended to guarantee convergence on a moral norm, and Benhabib just replaces (U) with two values that she thinks capture the 'meaning of normative validity.'

How does this discussion impress on Heath's views regarding nonstrategic bargaining, fair compromises, and normative validity? For Abizadeh, what is missing from Heath's analysis is some account of why agents should be *motivated* to accept a fair compromise in the first place. While Heath thinks that moral actors will accept the Nash bargaining solution on social norms "because it represents, roughly speaking, a fair compromise" (252), Abizadeh does not think this explanation is sufficient. Agreeing to the Nash solution requires more than a "commitment to the three axioms and to finding a compromise," it also

Requires a commitment to seeing what counts as a compromise through the lens of a particular and controversial theory of value. It requires each actor to see the *subjective* preferences of another as making some normative claim on him or her. (208)

As James Bohman puts it, "the formal theory of bargaining tells us very little, if anything, about the conditions under which compromises will be reasonable or acceptable to agents with normative orientations" ("Formal" 437). For instance, consider the third axiom (the IIA) in the Nash bargaining solution. The reason why the IIA is more controversial among game theorists than either the axiom of symmetry or efficiency is not only because there is empirical evidence that contradicts the consistency of preferences amongst bargainers but also because there are other "equally plausible" conventions that rational bargainers may or may not choose to follow and that will result in a solution other than Nash's.⁹ Heath does not give any reason why a rational bargainer should prefer the IIA over other candidates. Additionally, whatever alternatives count as 'relevant' or

'irrelevant' in reaching a fair compromise seem largely contingent on whatever value the agents assign to those alternatives in making their decision. Thus, in keeping with Abizadeh's contention that moral agents are motivated to accept compromises because they recognize that the preferences of others makes claims on them, it seems to be the case that there must be "principled as well as pragmatic reasons for embracing compromise" (Bellamy 114). How, then, do these two kinds of reason for accepting compromise find expression in complex and pluralist democratic societies?

3.3

There is consonance amongst almost all democratic theorists, even deliberative ones, that the voting mechanism will always remain the final arbiter for binding decision-making in complex democratic societies. At the end of the day, "democratic decisions are made by counting votes, not by assessing the rationality of the deliberation that precedes the votes" (Chambers, "Democratic Practices" 233). The consequence of the inescapably aggregative expression of individual wills articulated through majority voting, as opposed to a deliberatively considered common will, is that "some people must submit to an opinion that is different from theirs or to a decision that is contrary to their interest" (Przeworski 142). This is why many deliberative democrats, including Habermas, evoke at some point in their argument Dewey's edict:

Majority rule, just as majority rule, is as foolish as its critics charge it with being. But it is never *merely* majority rule... 'The means by which a majority is the more important thing': antecedent debates, modifications of views to meet the opinions of minorities... The essential need, in other words, is the improvement of the methods and conditions of debate, discussion and persuasion. (Facts and Norms 304)

Keeping this unavoidable constraint on deliberation in mind, I want to bring this thesis back to the question it started with, namely, can an analysis of compromise formation that stays consistent with the standards of Habermas's discursive concept of normative justification obviate or diminish in any way the force of certain criticisms leveled at his discourse-based democratic theory? It certainly seems that the possibility for compromises to be rationally grounded in a discourse diminishes the need to strongly "uncouple" opinion-formation in the public sphere from will-formation in parliamentary complexes and administrations, since one of the reasons Habermas does this is to preclude the public sphere from having to deal with strategic considerations that could potentially undermine its capacity to "perceive" and "thematize" social problems that parliaments themselves are incapable of detecting. One way to evaluate this question is to look at James Bohman's political conceptualization of a 'moral compromise' that puts certain features of our moral analysis of compromise formation into political practice.

Consider how from the perspective of rational choice theory, majority rule is simply a way of "externalizing" disagreement, a mechanism that is rationally acceptable since "when agents give up on argumentation and decide to vote, it can still be said they accept the final outcome for the same reason, because they are all motivated by the same inference from the same facts" (Heath 252). However, if democratic outcomes are such that they consistently do not reflect (or are at least *perceived* not to reflect) one's closest values, personal commitments, beliefs or integrity, the motivation for continued participation in such a scheme wanes and political apathy takes over. To avoid this state of affairs, one must find a means to make majoritarian democratic control more inclusive

and deliberative. As we will see, Bohman's idea of a "moral compromise" can be seen as incorporating inclusiveness and deliberation into the majoritarian procedure itself

Consistent with Abizadeh's insight that (U) represents the "meaning of normative validity," Bohman admits that there is a distinction between questions of what is good for some and what is just for all ("the evaluative self-understandings of persons or groups" versus "duties, norms and categorical imperatives")¹⁰ but expresses doubt that this distinction could do the work in political practice that Habermas believes it can. Habermas draws the distinction between moral, ethical and pragmatic discourses to give some idea of the warrants, based on the kind of question that is being posed, under which a claim to validity is acceptable. "Depending on the issue," Habermas asserts, "the various types of discourse and bargaining fill different roles for a rational political will-formation" (Facts and Norms 176). But not just any discussion, political or otherwise, is a discourse, for "conversations are more or less discursive to the extent that they approximate the *ideal* conditions of discourse" (Chambers, "Discourse" 234). Keeping this in mind, the problem with political deliberations, Bohman observes, is that they are hardly idealized, often the products of complex historical and social circumstances, and involve "various claims [that] are often mixed together so it is difficult to tell in advance what kind of reason will be convincing in any particular situation" (Bohman, Public 42). Thus, regardless of how carefully Habermas parses the distinction, in actual political deliberations it will often be the case that moral issues cannot readily be distinguished from ethical ones.

This observation leads Bohman to confront what he views as the implausibly stringent (and hence unhelpful in any practical application) conditions that define

Habermas's conception of public reasoning and of what kinds of claims constitute a public reason for him. For Habermas:

Anything valid should also be capable of public justification. Valid statements deserve the acceptance of everyone for the same reasons. The expression "agreement" is ambiguous in this respect. Whereas parties who negotiate a compromise might accept the result for different reasons, participants in argumentation must reach a rationally motivated agreement, if at all, for the same reasons. Such practices of justification depend on a *jointly and publicly reached* consensus. ("Reasonable" 86)

This definition of a public reason as one that can be accepted for the "same" reasons is

bound up with Habermas's strong regulating ideals of consensus and impartiality.

Bohman argues that *singular* conceptualizations of public reason such as these can have political purchase only at "the price of the plurality of potential preferences, norms, and values" (Public 81). Moreover, the convincingness of singular versions of public reason mainly relies on the capacity of participants to abstract themselves from their own particular historical attachments and traditional commitments. This requirement of public reasoning seems less practicable when one considers that liberal democratic societies are marked by competing, and in many ways, disparate views of what is good and what is just. Moreover, for a democratic proceduralism like Habermas's, where legitimate exercises of political power presuppose the command of the unanimous consent of those affected by political decisions, but which also rejects any "prior consensus on moral principles" or any concrete "concept of sovereignty...bound to the notion of embodiment in the people" (Facts and Norms 300), the singular account of public reasons potentially engenders an undesirable side effect for deliberative politics. For, if it must be presupposed that any public reason could potentially command unanimous consent but also be presupposed that there is no pre-existing harmony of interests among the people, "then [any dissenter] would seemingly be justified in questioning the very *legitimacy* of

all majoritarian outcomes, no matter how well-debated and revisable” (Rehg and Bohman 49).

To overcome this potential problem associated with the singularity of public reason presumed by discourse theory and the strong version of democratic legitimacy that accompanies it, Bohman proposes an alternative *pluralist* account of publicity, one that cannot only support a weaker version of democratic legitimacy but can also more easily accommodate the role of ‘reasonable compromises’ in cultivating heterogeneous but stable democratic societies. His proposal begins by making a clear distinction between these two types of validity:

On the one hand, public reason is singular if it represents itself as a singular norm of public deliberation; in light of this norm, agents come to agree upon some decision for the *same* publicly acceptable reason. On the other hand, public reason is plural if a single norm of reasonableness is not presupposed in deliberation; thus, agents can come to agreement with one another for *different* publicly acceptable reasons. (Bohman, Public 83)

As Heath’s earlier analysis showed, just because different people find different reasons convincing when they come to an agreement with each other about a common norm, it does not follow that they are acting coercively, duplicitously, surreptitiously or otherwise. Furthermore, the adjective ‘different,’ as it is used in the above passage, does not mean that simply *any* reason at all is publicly acceptable, for even “if we accept the *plurality of public reasons*, we need not *compromise on the normative syntax of public justification*”(Benhabib, Claims 143). ‘Syntax’ refers here “to certain structural features all statements that articulate would have to possess,” and Benhabib thinks public reasons count as public “because they could be defended as being in the best interest of all considered as equal moral and political beings” (140). Insofar as agents agree to

compromise for different reasons, those reasons still must meet a standard of public justification, and thus the plural potential of reasons does not discount their public nature.

Bohman uses this more pliable description of public reason to outline specific institutional arrangements and deliberative conditions that encourage continued participation and cooperation from citizens even when they disagree with the final decision. For, if unanimity on political matters with a moral significance is unlikely, it seems intuitive “that democratic institutions should not be constructed on the supposition that it is” (McCarthy 195). Alternatively, a more moderate view of democratic deliberation—one that expects “real conflicts that are so deep and persistent that no common framework for deliberation and reflection can be constructed without loss or coercion”—takes the position that in complex cases where deeply held values can be expected to come into conflict, democratic agreement should reflect the features of “a genuinely *moral* compromise in which plural public reason is exercised in the process of creating the framework for an ongoing public consensus, now a minimal one that demands only the willingness to continue to cooperate” (Public 84).

Democratic agreements based on a ‘genuine’ moral compromise are more demanding than compromises that either aim to maximize the overall aggregative satisfaction of individual preferences or assume a ‘prudential’ stance in which only minimal common ground for agreement is sought and contentious arguments over seemingly incommensurable values are avoided. A moral compromise is one that goes beyond bargaining for mutual advantage or “splitting the difference”—both of which often requires parties to either abandon some of their convictions (and usually some party more than another) or avoid talking about thorny and complex matters altogether—by

examining how particular conditions and structures for reaching minimal agreement today can “transform”¹¹ the existing framework for deliberation and interpretation and thus *expanding* the grounds for farther reaching agreements in the future. Rather than asking all involved to impartially distance themselves from their specific cultural commitments and to take the impersonal remains as the basis of agreement, moral compromises stress the importance of “recognizing the reasons of others and taking them up as parts of the overall solution” (93). On this interpretation of compromise, unanimous convergence on a particular agreement is not as important as creating further possibilities for deliberation: “a new moral framework, with new rules of cooperation and new forms of justification” (92).

By expanding the framework for deliberation and decision-making, genuine moral compromises aim to make it unreasonable for participants to walk away from those institutions altogether when they do not agree with the eventual agreement. *Unanimous* assent to the *outcomes* of political deliberation, as these outcomes are unlikely under conditions of value pluralism, is not what is important for a moral compromise; rather, what is needed is, as much as possible, unanimous support and acceptance for the institutionalized *procedures* in which the compromise was created. “To consider a political decision legitimate...it is enough [for citizen-deliberators] to assume that, given the conditions of deliberation, outcomes and decisions allow an *ongoing cooperation with others* of different minds that is at least *not unreasonable*” (Rehg and Bohman 46). But why do people continue to cooperate with each other, or participate in particular institutions, even when things do not go their way? Similarly, what is it about institutions that discourage people from continuing to cooperate in them? It is reasonable to assume

that people break off negotiations with others of different minds when they feel their opinions and beliefs are not being respected, and this result is not always avoidable through ostensibly impartial means of conflict resolution. Consider Bohman's claim that some institutions generally considered by most people to be impartial (like conventional majority voting procedures) are sometimes not perceived as so by groups that are on a politically unequal footing and find themselves occupying a "permanent minority status." If one is constantly being outvoted despite giving reasoned and thoughtful arguments for their position, it would be hard to make a case for why one ought to continue participating in this seemingly (from the perspective of those who always have their views confirmed by the process) impartial practice. Bohman thinks a democratic institutional arrangement reflects a genuinely fair moral compromise when it satisfies two conditions: if the arrangement takes "into account political inequality and they make possible continued participation of all groups in a common deliberative framework." (Public 96). He then offers several examples that meet these criteria, from modifications on traditional majority voting formats to make them more receptive to minority opinions and influence, to exemptions for certain religious or cultural groups who are obviously disadvantaged by the general intent of particular rules or legislation. This suggests that statements (like those of the authors introduced at the start of this section) about how majority decisions will always 'carry the day' when it comes to making democratic decisions need to be, in light of requirements of morality, subjected to further scrutiny. Additionally, these refinements do not subvert the essential principle that democracies built on ideals of "political equality" and "common citizenship," but "are based on the ongoing recognition of the need at times to modify the decision-making processes that

structure deliberation in order to make minority opinions more effective in securing public uptake” and “the ongoing process of public deliberation by which general norms are applied to specific situations” (98).

Bohman’s idea that moral compromises in the political sphere may well involve revising the procedures of decision-making and the structures of public deliberation seems to corroborate both Scheuerman’s and Young’s perception that Habermas’s conception of the political public sphere is too far removed from factual political authority and needs to be rethought. Nevertheless, I do not see how this change puts at risk the discourse theory as a normative basis for democratic legitimacy, nor do I think this change confirms Scheuerman’s belief that the relationship between communicative and administrative power is “inevitably destined to remain highly enigmatic” (Scheuerman 77). The mistake Scheuerman makes in his evaluation of Habermas is to think that deliberative ideals should (or can) be *directly* brought to bear on decisions taken within the administrative apparatus itself. This is a mistake because administrations, functionally speaking, are not concerned with the quality or inclusiveness of democratic deliberation and argumentation, but are programmed to use their organizational capacities so as to bring about the effective implementation of specific projects or the efficient allocation of resources. Moreover, large and heterogeneous democracies increasingly rely on these bureaucratic and administrative systems to ensure the orderly functioning of a society that might otherwise be overwhelmed by its own social complexity.

What the discussion of compromise has illustrated is that compromise formation and bargaining does not necessarily deform or damage the “structures of undamaged

intersubjectivity” that Habermas situates in political public spheres. Moreover, this suggests that public spheres, to retain an orientation toward “reaching understanding,” do not need to remain on the periphery of political decision-making processes to preserve their communicative potential.

For Bohman, while it is “precisely the demands of the public that give rise to the need for such administrative institutions,” the “execution of plans cannot be separated from public deliberation” (Public 189-91). What he means by this is that there are all sorts of strategies and forums for creating stronger, more inclusive, and more accountable interactions between the public and not just representative bodies but organized hierarchical bureaucracies as well. Bohman insists that all political institutions, regardless of how formal or how charged they are with being efficient, need to have pluralist public spheres take shape around them to make them publicly accountable and reflexive. This is especially true when issues of justice or fairness are involved. When formal bodies are forced to make concessions, tradeoffs, or compromises, or they choose one course of action over another for, perhaps, pragmatic, ethical, or moral reasons, their decisions still need to be explained openly to the public and be accountable to a more formal standard of democratic oversight. Among others forms, these public spheres might include public hearings, local meetings, citizen review boards and oversight committees, all of which serve “to develop relations of trust between administrators and public, not to create a counter-administrative set of institutions” (190). For, “without their own political public spheres, [bureaucratic and administrative institutions] are invested with too much unchecked power” (188). While this closer relationship between communicative and administrative power Bohman defines is, I suspect, still less clearly defined than

Scheuerman would like, it clearly depicts the potential for the political public sphere to more closely engage administrative bureaucracies, institutions, and “politics as usual” without forfeiting its status as a site of communicative freedom.

Conclusion

Compromises appear to be an inexorable element of democratic political life in complex and heterogeneous societies. Of course, not all compromises are created equal, and compromises have in the past (and are likely to continue to in the present) facilitated the continuation of injustice. However, genuine compromises must align with basic moral premises, and only then can expand the framework for continued cooperation amongst people whose ideas of the good and the just do not mesh together. Furthermore, compromises are not just about getting one’s fair share and they do not require a person to relinquish their deepest values or commitments, they only seem this way if we look at compromises solely in terms of the outcome rather than the negotiations that brought the compromise about in the first place.

It is pretty clear that Habermas confines compromise procedures to the “regulation of strategic interactions,” and this is why he considers it a “danger” if “compromise procedures [should ever] be applied to moral or ethical questions, so that these get *redefined* into strategic questions without anyone’s noticing or calling attention to the fact” (Facts and Norms 177). Yet, as my analysis indicates, genuine moral compromises presuppose a certain internal perspective on the part of discursively acting participants engaged in compromise formation, one that not only recognizes how the welfare of others makes a claim on their actions but also a willingness to transform their

own beliefs. The consequence of this interpretation of compromise formation for Habermas's discourse-based theory deliberative politics is not entirely clear; it is less than a 'democratization' of the administration but it also tempers the function of consensus in determining normative validity and democratic legitimacy. Also, this interpretation of normative compromise evidently makes certain demands of deliberating citizens that point beyond simply the discursive justification of fair procedures for compromise. Lamentably, since I have not developed any criteria here upon which to distinguish between those instances that call for a weaker reading of moral consensus and those that demand Habermas's stronger reading, the exact nature of the relationship between the universalization principle and moral compromise remains opaque. If I have done anything in this thesis, hopefully it has at least opened up some room for dialogue with those theorists who think Habermas's discourse-based explanation of democratic legitimacy is too strong to base a model of deliberative politics on.

Endnotes

¹ This passage is not found in the version of this “Three Normative Models of Democracy” that appears in Inclusion of the Other, but is in the version found in Democracy and Difference. Ed. Seyla Benhabib. Princeton: Princeton University Press, 1996. 21-30.

² Habermas’s political theory (and career) has endured a tendentious relationship to the social phenomena of mass protest and civil disobedience. I would point out that over the course of his career, in interviews and in text, Habermas has overlooked the correlation between protest and justice. In 1967, Habermas admonished German students protesting their government as “politically naïve,” and warned them that their actions could only fuel an emergent “left fascism”; in a 1986 interview, he spoke of how civil disobedience can serve a purpose in a ‘mature’ democracy so long as it both renounces violence, anchors its principles in constitutional essentials, and that it is a last resort... a particularly urgent appeal to the capacity and willingness for insight of the majority”; and, in Facts and Norms, he points to the *self-referential character* of civil disobedience, asserting that its purpose is to impel the government into “conflict mode,” and linking its justification to “a *dynamic understanding* of the constitution as an unfinished project.” While Habermas has over the years unquestionably amended his views of protest, I would argue that he too narrowly construes protest’s relation to the promotion of justice. And in tying civil disobedience so closely to constitutional essentials, essentials which can themselves be viewed as deeply context dependent (in that they are the product of the historical and social circumstances which gave rise to them), that he and Young end up talking past each other. To see this in the larger context of Habermas’s intellectual work, see Jordy Rocheleau “The Politics of Critical Theory: Discursive Proceduralism and Its Discontents,” Social Theory and Practice 29.1 (2003):137-57.

³ See Habermas, Facts and Norms 47-60.

⁴ “The practical discourses in which [claims on matters of justice] are examined will, then, have an internal relation to participants’ interpretations of their interests and hence to their values. (...) Questioning the very terms of debate, particularly with regard to the adequacy or appropriateness of standards of value, is the sort of thing that regularly occurs in the political discourse of pluralistic societies. In such cases, normative disagreements turn on value disagreements” (McCarthy, Ideals and Illusions 185).

⁵ Benhabib defines *universal moral respect* as the obligation “that we recognize the right of all beings capable of speech and action to be participants in the moral conversation,” and *egalitarian reciprocity* as the stipulation “that within such conversations each has the same symmetrical rights to various speech acts, to initiate new topics, to ask for reflection about the presuppositions of the conversations, etc” (“Afterword” 337).

⁶ Georgia Warnke neatly captures the basic idea of hermeneutic reflection when she writes that “the very ideas and assumptions that condition the attempt to construct universally valid principles of morality or justice are themselves those of a historical community.” Justice and Interpretation (Cambridge: MIT Press, 1993) 5.

⁷ The significance of a ‘fixed’ deadline for agreement is an important source of motivation for bargainers, for “if it did not matter *when* parties agreed, it would not matter whether they agreed at all” (Hargreaves and Hollis 135).

⁸ The meaning of the IIA is regularly depicted in the following example of its violation: A man sitting in a diner, having finished his lunch, asks the server for some pie and is given the option of cherry pie or apple pie. He chooses apple pie. The server retreats to the kitchen to get the pie, and returns a moment later to inform the man that blueberry pie is also available, to which the man replies, “Okay, I’ll have cherry then.” It should be noted that this example displays the IIA in its least controversial form. There are several versions that make choices like this one above more plausible.

⁹ Given that there are these other bargaining solutions that can be rationally chosen, this leaves open the possibility that some agents choose the Nash solution for strategic or instrumental reasons. One example of an alternative solution is associated with the Kalai-Smorodinsky *monotonicity* axiom.

¹⁰ Jürgen Habermas, "A Genealogical Analysis of the Cognitive Content of Morality," Inclusion of the Other: Studies in Political Theory, ed. Ciaran Cronin, Pablo De Grieff (Cambridge: MIT Press, 1998) 26.

¹¹ The idea of moral justification having a 'transformative' element is not explored in any detail in Habermas's account of the 'performative stance.' Politically speaking, Benhabib thinks a moral compromise "is a form of moral learning, ... [and is] not merely about arguments but about finding mutually acceptable ways of cooperating and continuing to exist with one another;" this seems of a piece with her earlier remarks that practical discourses necessarily have a 'moral transformative' moment insofar as that, by carrying out a discourse, "individuals come to realize a certain truth about their needs and interests and change their previously held beliefs about them" (Claims of Culture 145; "Critique" 312). I think an orientation toward "mutual understanding" implies that when we take up this stance in regard to someone else's ideas, interests and values, that we also take on a commitment to change or transform our own perspectives and preferences accordingly, especially if we believe that compromises are more than a *modus vivendi*, and that they are important for laying down foundations for future cooperation and further agreement in deeply pluralist societies.

Works Cited

- Abizadeh, Arash. "In Defence of the Universalization Principle in Discourse Ethics." The Philosophical Forum 36.2 (2005): 193-211.
- Baynes, Kenneth. "Introduction." Discourse and Democracy: Essays on Habermas's Between Facts and Norms. Ed. Rene Von Schomberg and Kenneth Baynes. Albany: SUNY Press, 2002. 1-30.
- . The Normative Grounds of Social Criticism: Kant, Rawls and Habermas. Albany: SUNY Press, 1992.
- Bellamy, Richard. Liberalism and Pluralism: Towards a Politics of Compromise. New York: Routledge, 1999).
- Benhabib, Seyla. "Afterword: Communicative Ethics and Current Controversies in Practical Philosophy." The Communicative Ethics Controversy. Ed. Seyla Benhabib and Fred Dallmayr. Cambridge: MIT Press, 1995. 330-69.
- . The Claims of Culture: Equality and Diversity in the Global Era. Princeton: Princeton University Press, 2002.
- . Critique, Norm and Utopia: A Study of the Foundations of Critical Theory. New York: Columbia University Press, 1986.
- Bohman, James. "Formal Theories, Pragmatic Purposes: Inferentialism, Rational Choice and Communicative Action." Canadian Journal of Philosophy 33.3 (2003): 423-40.
- . Public Deliberation. Cambridge: MIT Press, 1996.

-
- Chambers, Simone. "Discourse and Democratic Practices." The Cambridge Companion to Habermas. Ed. Stephen K. White. Cambridge: Cambridge University Press, 1995. 233-58.
- . "The Politics of Critical Theory." The Cambridge Companion to Critical Theory. Ed. Fred Rush. Cambridge: Cambridge University Press, 2004. 219-47.
- Cohen, Joshua. "Democracy and Liberty." Deliberative Democracy. Ed. Jon Elster. Cambridge: Cambridge University Press, 1998. 185-231.
- Couzens, David Hoy and Thomas McCarthy. Critical Theory. Cambridge: Blackwell, 1994.
- Gutmann, Amy and Dennis Thompson. Democracy and Disagreement. Cambridge: Harvard University Press, 1996.
- Jürgen Habermas. "Actions, Speech Acts, Linguistically Mediated Interactions and the Lifeworld." On the Pragmatic of Communication. Ed. Maeve Cooke. Cambridge: MIT Press, 1998. 215-55.
- . Between Facts and Norms. Trans. William Rehg. Cambridge: MIT Press, 1996.
- . "Discourse Ethics: Notes on a Program of Philosophical Justification." Moral Consciousness and Communicative Action. Trans. Christian Lenhardt and Shierry Weber Nicholsen. Cambridge: MIT Press, 1990. 43-115.
- . "Discourse Ethics, Law and *Sittlichkeit*." Autonomy and Solidarity: Interviews with Jürgen Habermas. Ed. Peter Dews. London: Verso, 1992. 245-72.
- . "A Genealogical Analysis of the Cognitive Content of Morality," Inclusion of the Other: Studies in Political Theory. Ed. Ciaran Cronin and Pablo De Grieff. Cambridge: MIT Press, 1998. 3-46.

-
- . Justification and Application: Remarks of Discourse Ethics. Trans. Ciaran Cronin. Cambridge: MIT Press, 1993.
- . "Popular Sovereignty as Procedure," Between Facts and Norms. Trans. William Rehg. Cambridge: MIT Press, 1996. 463-90.
- . "'Reasonable' versus 'True,' or the Morality of Worldviews." Inclusion of the Other: Studies in Political Theory. Ed. Ciaran Cronin and Pablo De Grieff. Cambridge: MIT Press, 1998.
- . "Reconciliation Through the Public Use of Reason." Inclusion of the Other: Studies in Political Theory. Ed. Ciaran Cronin and Pablo De Grieff. Cambridge: MIT Press, 1998. 49-74.
- . The Theory of Communicative Action (Vol. 2). Trans. Thomas McCarthy. Boston: Beacon Press, 1987.
- . "Three Normative Models of Democracy." Inclusion of the Other: Studies in Political Theory. Ed. Ciaran Cronin and Pablo De Grieff. Cambridge: MIT Press, 1998. 239-52.
- Hargreaves Heap, Shaun and Martin Hollis, Bruce Lyons, Robert Sugden, and Albert Weale. A Theory of Choice: A Critical Guide. Malden: Blackwell, 1992.
- Hargreaves Heap, Shaun P. and Yanis Varoufakis. Game Theory: A Critical Text. New York: Routledge, 2004.
- Heath, Joseph. Communicative Action and Rational Choice. Cambridge: MIT Press, 2001.

-
- Kellner, Douglas. "Habermas, the Public Sphere, and Democracy: A Critical Intervention." Perspectives on Habermas. Ed. Lewis Hahn. Chicago: Open Court, 2000. 259-87.
- McCarthy, Thomas. Ideals and Illusions: On Reconstruction and Deconstruction in Critical Theory. Cambridge: MIT Press, 1991.
- Munichs, Geert. "Rational Politics? An exploration of the Fruitfulness of the Discursive Concept of Democracy." Discourse and democracy : Essays on Habermas's Between Facts and norms. Ed. Rene Von Schomberg and Kenneth Baynes. Albany: SUNY Press, 2002. 185-99.
- Przeworski, Adam. "Deliberation and Ideological Domination." Deliberative Democracy. Ed. Jon Elster. Cambridge: Cambridge University Press, 1998. 140-60.
- Rehg, William. "Discourse Ethics." The Ethical. Ed. Edith Wyschogrod and Gerald P. McKenny. Malden: Blackwell, 2003. 83-100.
- . Insight and Solidarity: The Discourse Ethics of Jürgen Habermas. Berkeley: University of California Press, 1994.
- Rehg, William and James Bohman. "Discourse and Democracy: The Formal and Informal Bases of Legitimacy in *Between Facts and Norms*." Discourse and Democracy : Essays on Habermas's Between Facts and Norms. Ed. Rene Von Schomberg, Kenneth Baynes. Albany: SUNY Press, 2002. 31-60.
- Scheuerman, William E. "Between Radicalism and Resignation: Democratic Theory in Habermas's *Between Facts and Norms*." Discourse and Democracy: Essays on Habermas's Between Facts and Norms. Ed. Rene Von Schomberg and Kenneth Baynes. Albany: SUNY Press, 2002. 61-85.

Warnke, Georgia. Justice and Interpretation. Cambridge: MIT Press, 1993.

Young, Iris Marion. "Activist Challenges to Deliberative Democracy." Political Theory
29.5 (2001): 670-90.