



University
of Glasgow

Smith, Nicholas H. (1992) *Modernity, crisis and critique: an examination of rival philosophical conceptions in the work of Jurgen Habermas and Charles Taylor*. PhD thesis.

<http://theses.gla.ac.uk/1624/>

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

**MODERNITY, CRISIS AND CRITIQUE:
AN EXAMINATION OF RIVAL PHILOSOPHICAL
CONCEPTIONS IN THE WORK OF
JÜRGEN HABERMAS AND CHARLES TAYLOR**

by

NICHOLAS H. SMITH

A Thesis submitted for the degree of Ph.D. to
the Faculty of Arts, the University of Glasgow,
based upon research done at the Department of Philosophy

August 1992

© Nick Smith

ABSTRACT

In this thesis, I examine the rival conceptions of modernity, crisis and critique developed in the work of Jürgen Habermas and Charles Taylor. Since the publication of Habermas's highly influential *The Philosophical Discourse of Modernity* in the mid-1980s, scholarship on the conceptions of modernity and critique contained therein has gained its keenest focus in the context of the 'modernity vs. postmodernity' controversy. Meanwhile, in *Sources of the Self; the Making of the Modern Identity* - a book of comparable range and philosophical ambition to Habermas's study - Taylor has made his own distinctive contribution to what Habermas calls the philosophical discourse of modernity. But as yet, there has been no sustained investigation into the internal consistency and mutual challenge of the conceptions of modernity, crisis and critique defended by Habermas and Taylor. Taylor himself has recently proposed that a debate begin between what he terms cultural theory of modernity (to which his own work contributes), and acultural theory (to which Habermas owes allegiance). My thesis takes this invitation for debate as its point of departure for examining the competing claims of these two important philosophers.

The problem which organizes my contribution to a debate of the kind called for by Taylor is how, within the constraints of a philosophical conception of modernity, the claim to normativity can be brought to clarification. In chapter two, the sense in which the category of normativity is rendered problematic under conditions of modernity is explored. If the success of modern science shows that a moral order is no fit object of cognition, it can seem that the only rational action-orientation is instrumental in kind. I then introduce the strategies adopted by Taylor and Habermas for challenging this representation of the

modern tension between cognition and human identity. Chapters three and six are guided by the conviction that both Habermas and Taylor call upon a conception of crisis for the philosophical purpose of securing coherence to the idea of a moral order, and thereby the non-instrumentally rational redeemability of the claim to normativity. In chapters four and seven, the reflections upon language from which Habermas and Taylor derive conceptions of rational critique are examined. It is proposed that the distinctive significance each attributes to language can be understood in terms of the extent to which language features as a phenomenon to be grasped by cultural or acultural theoretical means. In chapter five it is maintained that Habermas's commitment to the latter informs a distinction between procedural practical rationality and therapeutic reason which is both difficult to sustain on its own terms and in tension with the work done by his conception of crisis. Chapter eight brings Habermas's and Taylor's alternative paradigms to subject-centred or instrumental reason into direct contact, and offers - albeit with a promissory proviso - a new way of comprehending the distinction between objective and subjective critical reflection. A hypothetical application of this distinction to the kind of critique appropriate to modernity's pressing ecological concerns concludes the thesis.

ACKNOWLEDGEMENTS

Thanks to Christopher Martin for support and faith. David Frisby and Shane O'Neill provided me with helpful comments on chapters four and five, for which I am grateful. I am particularly indebted to Irene Macleod for making the many typographical alterations to the final draft with an abundance of care, patience, and generosity.

CONTENTS

CHAPTER 1: Introduction	1
1.0 Foreword.	1
1.1 Two conceptions of modernity.	5
1.2 Two conceptions of crisis.	16
1.3 Two conceptions of critique.	23
CHAPTER 2: The Modern Tension between Cognition and Identity	30
2.0 Introduction.	30
2.1 Gellner on cognition vs. identity.	33
2.2 Morality and cognition - Habermas's response.	43
2.3 Morality and identity - Taylor's response.	51
2.4 Strategies for reconciliation.	59
2.5 Conclusion.	65
CHAPTER 3: Habermas's Conception of Crisis	67
3.0 Introduction.	67
3.1 Ethical totality and the causality of fate.	70
3.2 The fateful causality of psychopathologies.	81
3.3 Naturalistic objections.	94
3.4 Anti-naturalistic objections.	99
3.5 Conclusion.	103

CHAPTER 4: Habermas, Language, and Critique	105
4.0. Introduction.	105
4.1. Communicative and strategic linguistic interaction.	108
4.2. The primacy of communicative action; a bogus argument.	115
4.3. Consensus and strategy in language; two sets of objections	123
4.4. The primacy of communicative action; an alternative argument	133
4.5. Consensus and strategy in language; a third set of objections	137
4.6. Conclusion.	142
CHAPTER 5: Crisis and Critique; Tensions in Habermas's Conception	144
5.0. Introduction.	144
5.1. Diremption, differentiation and disharmony.	145
5.2. The divided labour of critical reflection.	151
5.3. Morality and ethical life.	156
5.4. Therapeutic and practical reason reconsidered.	176
5.5. Conclusion.	182
CHAPTER 6: Taylor's Conception of Crisis	187
6.0. Introduction.	187
6.1. Persons, strong evaluations, and the good.	189
6.2. The problem of self-reassurance reformulated.	196
6.3. The transcendental limit of human agency; identity crises.	204
6.4. Pragmatist objections.	211
6.5. Crisis and critical reflection.	221

CHAPTER 7: Taylor, Language, and Critique	230
7.0. Introduction.	230
7.1. The narrative form of practical reason.	232
7.2. Cultural theory and moral realism.	241
7.3. Language and the limits of cultural theory.	252
7.4. The expressivist mediation of morality and ethical life.	259
7.5. Conclusion.	266
CHAPTER 8: The Critique of Reason; Instrumental, Functional, or Procedural?	268
8.0. Introduction.	268
8.1 Procedural reason and performative self-contradiction.	271
8.2. The epiphanic mediation of cognition and identity.	281
8.3. The causality of fate again.	294
8.4. The objective and the subjective critic.	299
CONCLUSION: The Politics of Democratic Problem-Solving and World-Disclosure	304
Notes.	317
Bibliography.	354

CHAPTER ONE: INTRODUCTION

(1.0) Foreword

Marx once proposed that the task of critical philosophy is "the self-clarification of the struggles and the wishes of the age".¹ The proposal is at once challenging and disconcerting. For if philosophical critique is to bring to self-clarification the struggles and wishes of the age, it would need to have a concept of the age - a concept of modernity. But what would be *philosophical* about such a concept? And on what grounds could philosophers claim competence to participate in a 'discourse' of modernity?² These are troublesome questions, yet the idea that philosophical reflection can issue in some degree of self-clarification concerning the characteristic struggles and the wishes of the modern age informs contemporary philosophical debate of the very highest degree of rigour and sophistication. Much of the credit for this lies with the German philosopher, Jürgen Habermas. Habermas is the leading contemporary exponent of a tradition of thought called the 'Frankfurt School' of 'Critical Theory'. According to this tradition, pioneered in the 1930s by Max Horkheimer and Theodor Adorno, modernity is characterized by a domination of 'instrumental reason'. The instrumental society - the organization of the gamut of human affairs according to principles of technological control, anonymous administrative dictat, and profit-maximization - is, according to this school of thought, the paradigmatic location of struggle and discontent in the modern age. Indeed, the dominance of instrumental reason is so pervasive, the pioneers of Critical Theory maintained, that its critique requires an *epistemologically* distinctive, non-instrumentalizing form of philosophical reflection. Only if this requirement is taken seriously, they believed, could that which is dominated by instrumental reason be brought to self-clarification. In recent

years, historians of ideas have written extensively and illuminatingly on the continuity between Habermas's work and the founding ideas of the Frankfurt School.³ While I will occasionally draw from the results of these studies, I shall not directly be contributing to them.

However, recent research has also established that the Frankfurt School tradition by no means exhausts the possibilities for thinking about how philosophical critique of the kind called for by Marx might proceed. This literature can be broken down into two kinds of philosophical investigation. First, there are studies which offer readings of past philosophers - particularly Hegel and Heidegger - as responding to problems thrown up by a philosophical concept of the 'modern age'.⁴ Second, there is work which both establishes continuities between the motivating problems of Frankfurt School Critical Theory and rival contemporary frameworks for critique, and which assesses the relative merits of their proposed solutions. The main focus of attention here has been French post-structuralism.⁵ Habermas himself has defended his position against these rivals (especially the post-structuralists), which in turn has sparked off further research into the defensibility of his position.⁶ Although I will draw on some of this literature quite extensively, my contribution to the debate between Habermas and defenders of post-structuralism(s) will be - at best - tangential.

But this German-French axis of enquiry is currently undergoing a shift, due to the challenging proposals for a philosophically informed diagnosis of the 'spiritual situation of the age' offered by two English-speaking thinkers; Alasdair MacIntyre and Charles Taylor. Although both have engaged in controversy with post-structuralism, neither has yet

undertaken a sustained debate with Habermas's work.⁷ While a small amount of research has now been done by way of comparing the opposing positions of MacIntyre and Habermas concerning a philosophically informed normative critique of modernity, as yet there exists no sustained assessment of the apparently rival positions defended by Habermas and Taylor.⁸ The two rival conceptions developed by Habermas and Taylor of how best to undertake critical enquiry which aims at the "self-clarification of the struggles and wishes of the age" is the subject of this thesis.

My central focus of attention will be the sense in which "struggles and wishes" can stake or express a claim to *normativity*. The issue of how the category of normativity becomes problematic in modernity, and of how the category can be secured by appeal to reason, deeply structures the work of both Habermas and Taylor. It generates what Habermas calls 'the problem of modernity's self-reassurance', and contributes to what Taylor terms the 'crisis in confidence' in the 'modern identity'. Taylor agrees with Habermas that of the struggles of the age which call for self-clarification - the sources of moral/spiritual discontent which the concept of modernity covers - the dominance of instrumental reason is paradigmatic. Of the wishes of the age, both identify normative aspirations on the basis of which the dominance of instrumental reason can be overcome. They both offer a philosophically informed diagnosis of the struggle and wish for normativity which characterizes the modern age, but *how* their diagnoses are philosophically informed differs.

My primary objectives are to clarify and to assess the conceptual resources upon which Habermas and Taylor draw in their respective efforts at formulating and resolving, to use Habermas's expression, the problem of modernity's self-reassurance. In this opening chapter, I will introduce the philosophical conceptions of modernity, crisis, and critique

which will guide my reconstruction of the rival claims of these two thinkers. I begin (1.1) by drawing attention to Taylor's recent call for a debate between what he terms 'cultural' and 'acultural' theories of modernity. After explicating this distinction, and anticipating a problem which will later emerge in Taylor's defence of the 'cultural' position, I offer a schematic outline of Habermas's theory of modernity as a version of the 'acultural' kind. There are two obvious disadvantages to proceeding in this way: first, I must simplify Habermas's theory to an excessive degree; second, the introduction of technical terms which go insufficiently explained is unavoidable. On the first count, however, I think that the general strategy of Habermas's theory can fruitfully be reconstructed around certain 'acultural' claims, which I identify. On the second, the precise meaning of the Habermasian concepts which most interest us, such as 'communicative action' and the 'lifeworld', will be investigated in later chapters where I offer an interpretation of the role they play in Habermas's proposed resolution of the problem of modernity's self-reassurance.

My reading of both Habermas and Taylor centres around what I argue is the decisive role played by a concept of crisis in the normative foundations of their critical theories of modernity. In (1.2) I first introduce the formal characteristics of Habermas's crisis concept, and then outline how, according to Habermas, this concept enters into philosophical discourse as a problem of 'self-reassurance'. At this stage, I say little on how the concept of crisis structures the normative basis of Taylor's theory, except to anticipate my thesis - to be developed later - that there are tensions in the use Habermas puts to his conception of crisis which may be resolved by appeal to Taylor's conception. The latter, it will become clear, draws on figures of thought similar to those incorporated into the

'hermeneutic' philosophy of Hans-Georg Gadamer. Now the question of what kind of theory can best articulate the conception of crisis to which Habermas *himself* appeals is internally related to the issue of what it is in virtue of which a theory of modernity has critical powers. The latter issue has already been addressed in the well-known controversy between Habermas and Gadamer. In (1.3) I briefly outline the main contours of that debate, in order to lay the ground for an understanding of the divergent paths taken by Habermas and Taylor in pursuit of the goal of self-clarification concerning the aspiration and struggle for normativity characteristic of modernity.

(1.1) Two Conceptions of Modernity

In a recent article, Charles Taylor makes a distinction - and proposes that a debate begin - between "cultural" and "acultural" theories of modernity.⁹ By a "culture", Taylor means a practice expressing specific understandings of "personhood, social relations, states of mind/soul, goods and bads, virtues and vices, and the like".¹⁰ "To share a culture", to use Alasdair MacIntyre's rather more precise formulation of the same idea, "is to share schemata which are at one and the same time constitutive of and normative for intelligible action by myself and are also means for my interpretation of others".¹¹ The participants of a culture, accordingly, understand (are intelligible to) themselves through the schemata of conceptions of personhood, social relations, desirability-characterizations, and the like, which constitute the culture. The so-called "cultural" theories of modernity focus on the specifically modern schemata by way of contrasting them with equally specific others. The central issue for the cultural theorist, then, lies in identifying and explaining the transition from one schemata to another. Nietzsche, insofar as he portrayed modern scientific culture as one schemata of (ascetic) values in contrast and antagonistic to (life-affirming) others,

is cited as an exponent of this kind of approach. But Taylor suggests that cultural theorists have had little impact on the dominant canon of theory on modernity.¹² Taylor's purpose in initiating the debate is to give voice to what he perceives to be the unjustifiably silenced claims of a cultural conception and theory of modernity.

But Taylor acknowledges that this marginalization has not been completely without warrant. On the one hand, he suggests that practitioners of the cultural approach have typically been too one-sided in their identification of the schemata constitutive of modern culture (or as Taylor prefers to call it, for reasons I will explain later, the 'modern identity'). The content of the cultural conceptions of modernity have failed to do justice to the rich variety of, and inner tensions between, the modern schemata of self-interpretation. They have been too narrow, he claims, in their description of the aspirations characteristic of the modern age.¹³ Another reason for the lack of success of cultural theory, he implies, is the seemingly relativistic consequences of it; no appeal is made to a non-culturally specific and thus seemingly non-culturally relative ground for the critical assessment of practices informed by different shemata. The problem has been, Taylor suggests, that cultural theorists have been unable to *dispel* this appearance - they have failed to articulate convincingly why this relativistic conclusion does *not* follow. In other words, they have been unable to account for the idea that the transitions to the modern schemata may represent a gain or a loss in rationality. The very idea of a gain or loss in rationality appears to be problematic within the framework of a cultural theory, since what counts as rational is not independent of what a 'gain' or 'loss' means within the interpretive schemata of the theory itself. These gains or losses, if the cultural conception of modernity is to be preserved, cannot be theorized by appeal to a standard of rationality external to the

transitions themselves. *That* move, Taylor states, is distinctive of the "acultural" conception of modernity.

According to the so-called "acultural" theories, modernity is defined in terms of some "rational or social operation which is culture-neutral".¹⁴ By this is meant that modernity is theorized in terms of either the development of some *general capacity* for thought and action, or the performance of some social operation which is an independent variable of culture. In both cases, *all* cultures could, under suitable conditions, undergo the general transformation in terms of which modernity is conceptualized in the acultural theory. Any particular culture could serve as "input" to the general operation, and the operation is definable independently of any specific culture. General capacities for thought and action typically invoked by theories of this kind, Taylor observes, include the "scientization of world-view", the capacity for discrimination between "fact and value", and the ability to perform actions which bring about an end by the most efficient means. Of the culture-neutral social operations taken to characterize modernity, Taylor mentions increased mobility, urban demographic concentration, and industrialization. Theorists who adopt this kind of acultural approach may either affirm or deny that modern societies are successful in their performance of the selected function. For instance, there are those who affirm that modernity has successfully fostered an objective cognitive grasp of nature, that the growth of scientific consciousness in modern societies actualizes the capacity for separating fact from value, and that with modern technology, the ability to bring about an end by the most efficient means reaches unprecedented fruition. On the other hand, there are those who deny that modernity does satisfy the aculturally defined requirement. Thus, in the case of the social changes characteristic of modernity, they deny that the chosen culture-neutral

operation is successfully performed. The secure social integration of individuals, for instance, might be cited as a general operation which pre-modern, tightly-knit, relatively immobile, traditional communities do best.¹⁵ The point is that any culture could in principle serve as input for the capacity to view the world scientifically, or to live in organic solidarity with others. Another way of putting Taylor's point, would be to say that the standards against which modernity is judged are not taken from any particular culture, but that all cultures are in principle accountable to these general, culturally non-specific standards or capacities.¹⁶

In the article in which Taylor expresses the wish to initiate a debate between cultural and acultural theories of modernity, his official view is to deny that "one can make an exclusive choice between them".¹⁷ Cultural theories of the kind he wants to defend, he acknowledges, neglect certain crucial facets of the transformation to modernity - for instance, that "modern science *has* a validity, and the accompanying technology an efficacy, that we have 'come to see'". From this statement, it seems that while not wanting to dismiss the claims of acultural theories, Taylor simply seeks to redress the balance between them and cultural theories. Unless the cultural approach is given its due, Taylor asserts, we are in danger of either the ethnocentrism of misclassifying what is specific to modern western cultures as a universal given (because it is construed by the acultural theory as culture-neutral), or of misunderstanding the various dimensions of just what it is which is specific to modern western cultures (because the concepts of an acultural theory are not geared towards showing the contrast between modern and other cultures). But as we will see, the position Taylor defends elsewhere is actually stronger than this. For the *sense* in which science has a validity, he will argue, can be brought to clarification by appeal to the

resources which only a theory of the *cultural* kind makes available. The *means* by which we "come to see" the validity of science, he claims elsewhere, is *no different* to the means by which we come to see the validity of schemata of self-interpretation. In chapter seven, I will argue that there is an unresolved tension in Taylor's defence of a cultural theory of modernity which corresponds to this oscillation between the weak claim of redressing the balance between cultural and acultural theories, and the strong claim that only a cultural theory can articulate the transformations to be explained.

Having expressed the desire to initiate a debate between cultural and acultural theories in general, Taylor makes only one explicit reference to the sense in which Habermas's theory of modernity in particular is acultural in type.¹⁸ In the remainder of this section, I want to introduce Habermas's conception of modernity as part of an acultural theory. The justification for so classifying Habermas's theory of modernity is apparent from the basic insight around which it is fashioned; that since the seventeenth century, distinctive structures of thought and action have evolved within western societies, and that this evolution can (and must) be conceived as a process of rationalization.¹⁹ The essence of Habermas's conception is that modernity stands in an "internal relationship" to rationality,²⁰ and the central task of his theory is to establish the precise nature of this relationship. The strategy he adopts in undertaking this task can be informatively reconstructed, I believe, around the following guiding claims: (1) that there are certain 'resources' which all societies must make available if they are to be capable of reproducing themselves 'non-pathologically'; (2) that it is possible to reconstruct ideal conditions the satisfaction of which would render the availability of such resources 'rational'; (3) that modern societies can accurately be described as approaching this ideal type; (4) that allegiance to the 'project

of modernity' involves a normative commitment to the closing of the gap between actually existing modern societies and the ideally projected rational type; and (5) that such normative commitment is itself rationally justifiable. At this point, my sole aim is to indicate how a certain conception of modernity emerges from this *kind* of strategy - one which, according to Taylor's nomenclature, is acultural in orientation. Only later will I analyse how Habermas attempts to justify the decisive claims on which this strategy rests.

(1) Habermas makes a distinction between the 'symbolic' and the 'material' resources which societies must make available if they are to sustain their identities as collectivities. Of the symbolic resources, he distinguishes what he calls 'meaning', 'solidarity', and 'personal identity'. 'Meaning' is a term of art which designates the resource which is provided by the stock of largely implicit, historically generated knowledge upon which participants in everyday communication draw; it is that in virtue of which social actors are capable of engaging in intelligible action. 'Solidarity' refers to that in virtue of which individual actors are bound together as a community - it covers obligations, normative prescriptions for action, and nourishes a sense of communal belonging. Concerning meaning and solidarity, there must be some degree of consensus if the 'symbolic' identity of the society is to be maintained. 'Personal identity' is what enables autonomous action on the part of individuals - it provides that sense of individuated selfhood in virtue of which individuals can direct, and be responsible for, their own actions. Habermas insists that each of these resources, if they are to be capable of maintaining the collective identity of the society, must feed off the power of conviction. To provide them is the job of three different social mechanisms; what Habermas calls 'cultural reproduction', 'social integration', and 'socialization', respectively. Through cultural reproduction, inherited stocks of knowledge

and interpretive schema are passed from one generation to another; through social integration, actions are coordinated on the basis of intersubjectively recognized norms; and through socialization, individuals learn to differentiate themselves from others and to become accountable for own actions. The integrity of the society depends on each of these mechanisms being performed satisfactorily - on making these core symbolic resources available. But besides these symbolic resources, of course, any society must also make certain 'material' resources available - such as food, shelter, energy, and the like. To make these resources available, is for the society to generate its material means of subsistence.

(2) Habermas contends that it is possible to reconstruct how these socially necessary mechanisms can operate 'rationally'. Again, Habermas distinguishes two different forms of rationality appropriate to the reproduction of the two different kinds of resource. The symbolic resources are reproduced rationally if they are subject to agreements between participants in dialogue reached solely on the basis of the better argument. The mechanism here is a formally definable procedure of linguistic interaction which issues in a consensus reached with minimal appeal to some pre-given, time-honoured authority. A 'culture' reproduced in this way would take on an increasingly reflexive character, as inherited stocks of knowledge are subjected to a continuous process of revision on the basis of their ability to resist criticisms of their validity. 'Solidarity' would be generated according to principles which were capable of passing the legitimating test of consensus reached through argument. The capacity of action-coordinating norms to carry conviction would be displaced from the concrete content of particular normative claims to the formal procedure of justifying them. Such principles would thereby take on a generalized, universal character. And without the security of guaranteed stocks of knowledge and fixed patterns

of solidarity to rely on, the pressure increases towards highly individuated personal identities; abstract and fragile 'ego-identities' which must direct their own life-projects. Habermas calls the vehicle through which the symbolic resources are made available rationally 'communicative action'. The mechanism through which material resources are provided for rationally, on the other hand, is the efficient control over, and adaptation to, an environment. And this requires, Habermas maintains, that the mechanism through which actions are co-ordinated for the purpose of renewing the material resources of society becomes separated from the mechanism for rationally coordinating actions serving to reproduce the symbolic resources.

(3) This separation is decisive for Habermas's conceptualization of the specific kinds of crisis to which modern societies are disposed. They are consequences of what Habermas calls 'societal rationalization', the main contours of which I have so far been summarizing. But supplementing his conception of modernity as rationalized along these lines, is a characterization of modernity as a rationalized 'culture'.²¹ I shall refer to this dimension of rationalization to indicate how the actual transformation to modernity can be seen to fit Habermas's rational projection.

The defining characteristic of the products of cultural modernity, in Habermas's conception, is that they can each be compartmentalized as contributing to one of three distinct spheres of 'value'; science, moral understanding, and art.²² The value of any cultural artefact, under conditions of cultural modernity, depends on its contribution to any *one* of these spheres. Accordingly, cultural modernity dictates that any of its products be assessed either *qua* scientific theory, or *qua* theory of morality, or *qua* work of art (or art criticism). In

the institutionalized embodiment of each sphere, the specific kind of claim immanent to the artefact is rendered autonomous, and thematized in a manner which allows for a specialized, expert competence in matters scientific, moral, and aesthetic. The 'culture' of cultural modernity is, in Habermas's specific sense of a 'rationalized culture', one of experts in different kinds of discourse and judgement. Thus scientific discourse is institutionalized in such a way as to thematize the kind of validity claim immanent to scientific theories, independent of the contribution such theories might make to the understanding of morality, or in virtue of their merit as works of art. Similarly, questions of justice and morality require their own independent and specialized treatment, while the production and criticism of works of art takes on a validity which is independent of their merit as scientific or moral claims, and for which a quite separate kind of specialized competence is required. For Habermas, the value of a scientific claim, *qua item of cultural modernity*, lies in its truth; truth is the aspect of validity which is thematized in scientific discourse. The value of moral discourse, on the other hand, resides in how well it thematizes problems under the validity aspect of justice or normative rightness. The value of modern art and art criticism lies neither in its truth nor in its normative rightness, but in its exploration of subjective authenticity and beauty.²³ According to Habermas, then, 'cultural modernity' refers to the separation within culture of science, morality, and art; to the thematization within science, morality and art of claims to truth, normative rightness, and subjective authenticity respectively; and to the separation between the institutions through which this thematization issues in 'expert cultures', from everyday life.

(4) Habermas traces the embryonic normative commitment to a form of life rationalized along these lines back to the Enlightenment. The 'project of modernity' begins with the

Enlightenment insistence upon *the failure of religion to furnish the resources outlined above*. At the level of culture, religion abrogates the differentiation between science, morality, and art, which to the thinkers of the Enlightenment, had proved itself as the unassailable condition of reflective, self-conscious knowledge. Feeding off the self-sanctioned knowledge of authority and tradition, it seemed, religion hindered that shift towards reflexivity which conditions the aspiration towards genuinely valid knowledge claims. At the level of solidarity, it imposed exemplary hierarchical models of the ordering of society from a past bereft of legitimating weight, hindering that shift towards genuinely universal norms which reflection demands. And at the level of personal identity, religion had defined a place for each individual in this social order, hindering the shift towards self-directed autonomous action. The Enlightenment critique of religion (and more generally, of authority and tradition), then, was informed by the ideal of a rational practice in which the conditions for "self-consciousness" (reflexively valid knowledge), "self-determination" (reflexively valid norms), and "self-realization" (of a reflexive, highly abstract personal identity) are established. Without an exemplary rational order drawn from the past or the cosmos on which to model itself, the project of modernity must generate conviction concerning its norms, as Habermas puts it, "out of itself".²⁴ Since Habermas assumes that the means of rationally coordinating actions which reproduce material resources are in themselves norm-free, it follows that the modern project cannot but insist that its norms be generated *out of the mechanism for rationally reproducing its symbolic resources*.

(5) It is by following the strategy just outlined that Habermas theorizes the transition to modernity as a process of rationalization. The theory explains this process according to

"internally reconstructible sequences of stages of competence" ²⁵ in the reproduction of both symbolic and material resources. Its 'acultural' orientation should now be evident. The first stage proposed that *all* societies, irrespective of their particular culture, must reproduce symbolic and material resources according to an "internally reconstructible" mechanism. This is taken to hold independently of the content of the resources themselves. In Habermas's terminology, "the lifeworld" (considered as a resource for self-interpretation) displays a *structural* invariance which underlies the diversity of particular cultural contents; "particular forms of life, which emerge only in the plural,...exhibit structures common to lifeworlds in general".²⁶ The second stage outlined two culture-neutral operations which would process any 'input' rationally; consensus reached through argument and efficient adaptation to an environment. Habermas also maintains, as I mentioned, that it is possible to internally reconstruct competence in each in terms of their differentiation. As is clear from the third stage, differentiation within 'culture' conditions competence in reaching rational agreements about truths and norms.

But these differentiations make the rational justification of allegiance to the project outlined in the fourth stage problematic. We saw that, for Habermas, only that kind of competence which is required for rationally reproducing symbolic resources has normative relevance. But is a society which reproduces its symbolic and material resources by these rational mechanisms capable of sustaining (and stabilizing) itself? Are rationally justifiable principles available which can integrate the actions of abstract, highly individuated self-directing actors? Most generally, can the struggles and aspirations of modernity be brought to self-clarification from *within* the project of modernity so understood? In the light of these questions, a need for self-reassurance emerges on the part of the project of modernity.

(1.2) Two Conceptions of Crisis

Internally related to the two conceptions of modernity developed by Taylor and Habermas, are conceptualizations of 'crisis' to which modern individual and collective identities are disposed. Being so disposed, modernity stands in need of 'self-reassurance'. In this section, I will first introduce the formal attributes of Habermas's concept of crisis, before indicating how a problem of self-reassurance can be articulated in their terms. Central to my thesis will be the claim that *in addressing the problem of self-reassurance*, Habermas appeals to a model of crisis which is incompatible with his conception of modernity *insofar as it is part of an acultural theory*. I will also claim that this incongruity can be rectified by a modified appeal to the conceptual resources Taylor's preferred cultural approach makes available - resources which are implicitly tapped by Habermas himself.

On his way towards specifying a 'social-scientific' concept of crisis - one which is suitable for an acultural theory of modernity - Habermas mentions two other contexts in which the concept of crisis has application; the medical and the tragic.²⁷ In the medical context, crisis refers to "the phase of an illness in which it is decided whether or not the organism's self-healing powers are sufficient for recovery".²⁸ The person's critical illness, his or her body's externally caused deviation from its normal healthy state, is an objective reality upon which the life or death of the person turns. From the point of view of the physician, there will be empirical criteria for determining this malfunction as an objective process. But the peculiarity of a crisis, Habermas notes, is that it "cannot be separated from the viewpoint of those undergoing it". The illness takes the form of a crisis insofar as the person is "a subject condemned to passivity".²⁹ In the medical context, therefore, the concept of crisis is associated with "the idea of an objective force that deprives a subject of some part of his

normal sovereignty. To conceive of a process as a crisis, Habermas continues, "is tacitly to give it a normative meaning - the resolution of a crisis effects a liberation of the subject caught up in it".³⁰

According to this characterization, crises are (1) inseparable from the point of view of those suffering them, (2) caused by influences external to a subject which appear as something objective, (3) normatively ascribed by (often implicit) appeal to a criterion of emancipation, and (4) resolved with the effect of emancipation. The distinctive twist which Habermas will give to these attributes is anticipated by the second context he chooses to consider for the sake of clarifying the crisis concept. Following Hegel's understanding of crises as they feature in classical tragedy, Habermas remarks that;

crisis signifies the turning point of a fateful process that, despite all objectivity, does not simply impose itself from the outside and does not remain external to the identity of the persons caught up in it. The contradiction, expressed in the catastrophic culmination of conflict, is inherent in the structure of the action system and in the personality systems of the principle characters. Fate is fulfilled in the revelation of conflicting norms against which the identities of the participants shatter, unless they are able to summon up the strength to win back their freedom by shattering the mythical power of fate through the formation of new identities.³¹

Although Habermas immediately distances himself from the millenarian overtones of the philosophy of history implicit in this conception - one which he aims to replace with a scientific reconstruction of the logic of learning processes (an "internal reconstruction of stages of competence") marking the transition to modernity - he will continue to rely on key aspects of this tragic conception of crisis when articulating what he takes to be the driving philosophical problematic of modernity's self-reassurance.³² Before turning to that problem,

let me briefly clarify what these features are by comparing them with the medical usage of the crisis concept. As in the medical context, where the objectivity of the critical condition from the outside point of view is suffered as a crisis only from the inside perspective of the subject of it, so here the effect on the internal identity of subjects of the objective unfolding of the dramatic events makes for the crisis. The crisis occurs due to circumstances which are external to and out the control of the particular characters, yet they come to inflect the identities of the characters themselves. And again like the medical model, the effect of a successful resolution of the crisis is liberation - the characters "win back their freedom". But despite these similarities, there are differences which are not simply due to the greater sophistication of the tragic model. The unfolding of the plot is not 'objective' in the same sense as the unfolding of an illness is, and the outside point of view of the gods, the story-teller, or the audience, is not the *same* outside point of view as of the physician. But more to the point, whereas in the medical model the objective process which plunges a person into crisis is a physical event of nature (a disease), in the tragic model it is human actions and personalities which are ultimately responsible for the crisis. To be sure, these actions and personalities have an 'inherent' structure, but it is one which remains behind the backs of the participants until it is revealed *in* the conflict, bound to catch up with them, which plunges their identity and existence into crisis. In the tragic model, the crisis has a causality of 'fate' rather than nature.

I now want to indicate how Habermas draws on this conception of crisis to articulate how the problem of modernity's self-reassurance enters into philosophical discourse. As we saw in the previous section, the project of modernity, as initiated by the philosophers of the Enlightenment, commits itself to the norms of a 'rational practice'. Such norms are

'unconditioned' (rational) by virtue of holding independently of historically contingent traditions. Modernity must generate its norms "out of itself", which means - if my reconstruction of the acultural strategy of Habermas's theory is correct - out of the mechanism for reproducing its symbolic resources. The project of Enlightenment, Habermas informs us, emerged as a reaction to a degenerate form of religion as the provider of symbolic resources, insofar as it subordinated the individual's reason to time-honoured authority, and proved scarce in resource for nourishing solidarity between individuals. Religion, as it appeared to the philosophers of the Enlightenment, failed to provide crucial symbolic resources which - in the last resort - can sustain an identity only by virtue of their *power to carry conviction*. Habermas then describes how the Enlightenment philosophers turned instead to self-consciousness, self-determination, and self-realization, as the orienting norms of the project of modernity. These norms, it was thought, could be conceived as manifestations of the "principle of subjectivity", which - according to Hegel - is "the principle of the modern world" itself.³³

But Hegel also perceived, Habermas reminds us, that the Enlightenment critique of religion, performed in the name of reflective reason and the sovereignty of the rational subject, is *itself* incapable of furnishing the symbolic resources necessary for sustaining the identities of individuals and collectives. "Subject-centred" reason, or reason as it is articulated in the principle of subjectivity, leaves itself without a motivating power for the individual who must act, and without a unifying power for the collective which must mediate individuals. Incapable of "interesting the heart and of having an influence upon feelings and needs",³⁴ the principle of subjectivity merely perpetuates the failures of the principles of degenerate religion it was to replace. For Hegel, categorial oppositions central to the principle of

subjectivity - such as between "nature and spirit", "theoretical and practical reason", "I and non-I", "knowledge and faith" - reflect the historically rooted instabilities of the modern world. This instability Hegel attributed to a "sundered harmony of life", to real "diremptions" emerging in modern society.³⁵

Hegel also refers to the abstract oppositions of the principle of subjectivity as expressed in enlightenment thought as reflections of a concrete "estrangement of spirit" in modern conditions of life - an estrangement which he captures in the notion of "positivity". For Hegel, Habermas tells us, positivity represents both the "signature" and the "need" of an age embodying subject-centred reason. The positivity of reason, like the positivity of religion, refers to a withdrawal of motivational and unifying power only now from the very norms which apparently give the project of modernity its orientation. It refers, in other words, to the failure of both religion and Enlightenment to provide sustainable - because capable of carrying conviction - symbolic resources. Hence a philosophical problem of self-reassurance faces the project of modernity. Self-reassurance is required because, as Hegel conceived it, a form of life oriented by the norms contained in the principle of subjectivity leads to experiential (personal) and public (institutional) consequences which simultaneously undermine the stability of that life form. Hegel's critical intuition, then, concerns the modern identity's intrinsic tendency, or fateful disposition, towards crisis.

Habermas identifies this theme of diremption as the underlying problematic of the philosophical discourse of modernity; he designates it as the source of the need for modernity's self-reassurance.³⁶ How, then, is the positivism of reason to be theorized given the conception of modernity introduced in the previous section? And how is the crisis

which it represents to be conceptualized in a social scientific discourse?

I noted in the previous section that, according to Habermas's theory, modernity is characterized by a separation of the mechanisms for reproducing its symbolic and material resources. This separation, he argues (on the basis of an internally reconstructible sequence of stages of competence), allows both to become rationalized - but the meaning of rationalization differs in the two cases. This difference, he maintains, must be taken into account in the methodological orientation of the theorist. Society as a rationally organized 'system', according to Habermas's usage, is theoretically comprehensible from the external point of view of an observer as a self-managing functional organism. The rationality of a system is measured in terms of its functional efficiency in self-preservation. This it achieves by way of a growth in complexity and material production. System-maintenance depends upon maximally efficient integration of action consequences, and this is achieved by what Habermas, following Talcott Parsons, calls 'steering media'. A society's symbolic resources, on the other hand, being the reservoir of meanings, can only be grasped from the theoretical perspective of a participant in communication. Considered as this resource, what Habermas calls the lifeworld becomes rationalized to the extent that it is reproduced by communicative participants in rational dialogue with each other. Habermas's thesis is that modernity unfolds initially as an uncoupling of lifeworld from system, and then degenerates, under the pressures of an expanding system, into a state of colonization of the lifeworld by the system. Within the functional subsystems of the capitalist economy and the modern bureaucratic state, action integration is mediated not communicatively, but by the 'delinguistified' steering media of money and power. These react back and mediatize or instrumentalize the symbolic reproduction of the lifeworld, and hence put the

identity-forming process of individuals and groups in jeopardy. There comes a point when the integrity of the lifeworld is threatened by systemic mediatization. At this crisis point, the economic/bureaucratic system colonizes the lifeworld. The result is a pathological systemically induced 'reification' of the lifeworld. This manifests itself to the participants as a 'one-sided rationalization' of everyday life around the 'instrumental' dimension of rationality which is exploited in system expansion. This, according to Habermas, is how the positivism of reason which has motivated the philosophical discourse of modernity since Hegel, can be conceptualized without resort to the principle of subjectivity and subject-centred reason.

The two theoretical attitudes which Habermas holds to be required for understanding the paradox of rationalization - namely the undermining effects of a rationalized system on a lifeworld whose rationalization it presupposes - imply two different conceptions of crisis. 'System-crises' operate behind the backs of the participants and can be theorized only from the vantage point of the first of the two methodological orientations mentioned above; 'lived crises' manifest themselves *in* the lifeworld and are caused by deformations *of* it.³⁷ While it would be a considerable virtue of Habermas's theory if it were able to establish the interconnection of these two conceptions, his strategy runs the risk of displacing the theoretical perspective on the lifeworld from the internalized, action-orienting perspective of the agent: can Habermas convince us that the lived or existential crises which he wants to explain - those which react to the modern dominance of instrumental reason - are best articulated in a theory with conceptual resources designed to account for the above paradox of rationalization? Is the paradox of rationalization - put as it must be in acultural terms - an adequate basis for bringing the modern struggle for 'normativity' to self-clarification?

In the next section, I turn to a debate which prefigures Habermas's and Taylor's opposing positions on these questions, in that it addressed itself to the philosophical basis of the claim to normativity.

(1.3) Two Conceptions of Critique

Central to any debate between acultural and cultural theories of modernity, must be the question of how adequately each is able to articulate that in virtue of which it possesses *critical* powers. As theories, both contain commitments to the view that this adequacy can be defended through rational reflection. Given this shared commitment, the failure of either kind of theory to avail itself of conceptual resources required for making intelligible its own possibilities and potentialities for critique must count as a *prima facie* case against that kind of theory. Since Taylor's understanding of the status of cultural theories is avowedly indebted to Gadamer's conception of hermeneutics,³⁸ the much-discussed controversy between Habermas and Gadamer offers itself as a useful point of departure for grasping the differences between the foundations for a normative critique of modernity proposed by Habermas and Taylor. After outlining the main contours of the earlier controversy, I will indicate how a debate between cultural and acultural *critical* theories of modernity of the kind called for by Taylor follows naturally from and marks an advance upon the former debate.

The controversy between Gadamer and Habermas centres around a constellation of claims concerning the scope and function of 'hermeneutic reflection'.³⁹ Hermeneutic reflection is, in the first instance, reflection upon what it is in virtue of which an interpreter is capable of reaching an understanding of an initially unfamiliar (because historically or culturally

distant) text. The first principle which hermeneutic reflection reveals, according to Gadamer, is that the interpreter cannot help but bring to the text anticipations of its meaning; anticipations which are not the interpreter's own invention. The interpreter does not suddenly appear before the text as a *tabula rasa* - the text is always approached *from* somewhere. Gadamer chooses to call this 'somewhere' "tradition", and he designates these anticipations "prejudices".⁴⁰ These pre-reflective prejudices, and the tradition which carries the interpreter to the text, are the interpreter's *access* to it. It is mistaken, therefore, to think that prejudice and tradition are something which merely cannot be *avoided*; rather they are a positive condition of the interpreter's possibility of reaching an understanding. Further, the text itself is also both the bearer of and carried by tradition. Consequently, the point of departure for hermeneutic reflection is the concrete historical positioning of interpreter and interpreted; what Gadamer calls the "hermeneutic situation". The interpreter is always situated in his or her attempt to reach an understanding, a task which is only intelligible by virtue of the prejudices which are shaped by a tradition within which both interpreter and interpreted always find themselves.

A second principle which hermeneutic reflection reveals is that the understanding which is sought on the part of the interpreter is reached through a procedure of dialogue with the text which is interpreted. The kind of understanding which the interpreter seeks is "dialogical" in that it involves the reaching of an agreement, with the person who speaks through the text, concerning the subject-matter of the text. But in order to avoid misconstruing the nature of this agreement as the coincidence of self-transparent psychological contents, and to give due weight to the traditions and prejudices along which subjectivity is always carried, Gadamer coins the phrase "fusion of horizons" to describe

the phenomenon of reaching a common accord.⁴¹ In this fusion of horizons, the truth of the subject-matter about which understanding is sought discloses itself. Further, it is the function of the language of interpretation to disclose such truth. Interpretation, then, aims at truth which is disclosed through a fusion of the horizons between interpreter and interpreted, and this results from a genuinely dialogical interaction between the carriers of tradition. Consequently, the goal of interpretation is properly conceived as the broadening of the horizon of the interpreter, and thereby an enriched self-understanding. And this is a third and crucial principle brought to hermeneutic reflection; that the understanding which is sought has a *productive*, practical character. As well as being pushed *from* somewhere - the anticipations and prejudices which inform and guide a tradition - the interpreter is also pulled *towards* an expanded horizon which cannot be anticipated prior to a dialogical interaction with the text. And through this process of interpretation, as it is revealed to hermeneutic reflection, both interpreter and interpreted are mutually transformed in a non-arbitrary, practically efficacious, truth-disclosive manner.

Now Gadamer holds that the position of the interpreter - the hermeneutic situation - is paradigmatic for the understanding of human interaction as such. Consequently, in its scope hermeneutic reflection has a claim to "universality". All human understanding, Gadamer insists, contains a substratum of prejudice which resists reflective rationalization. Gadamer proposes a general "rehabilitation" of the concepts of prejudice, tradition, and authority, which he thinks have been negatively polarized against an abstract, ahistorical "Enlightenment" conception of reason. According to the Enlightenment model, as Gadamer interprets it, rational thought and action is *defined* in opposition to the recognition of authority and tradition. But this ideal of rationality, Gadamer claims, breaks with the

principle revealed by hermeneutic reflection that understanding only issues from within or between tradition(s) - and that without prejudices, the human enquirer would be without any "windows" to the world. A consequence of the contextual, historical character of understanding, is that there can be no "Archimedean point" - to use Descartes' notorious metaphor⁴² - independent of the content of tradition, from which to assess the rationality of prejudices. Likewise, if understanding human thought and action proceeds in the way Gadamer describes, then a method which orients itself to neutralizing prejudices - such as the one employed in the natural sciences - is fundamentally inappropriate for grasping the significance of human affairs. Conversely, if it is accepted that the goal of understanding is a truth of some sort, then truth itself needs to be divorced from a method which, if followed, would guarantee it "monologically". There can be no such guarantee if the means by which understanding is reached is dialogical, for there can be no telling in advance what the outcome of the dialogue, conducted through language, will be. Indeed, the scope of hermeneutic reflection is none other than the scope of language itself, and the scope of language is universal; it covers all meaningful human activity in its "world-disclosive" aspect.

It is this last claim which provokes Habermas's challenge. While he finds much that is acceptable in Gadamer's hermeneutic insights - particularly his emphasis on the linguistic, participatory, dialogical character of reaching an understanding, and the limits of grasping the meaning of social action from the point of view of a scientific observer - he is unable to accept the contention that hermeneutic reflection is universal in its scope. Habermas's worry is that by ascribing universality to hermeneutic reflection, Gadamer forfeits the critical potentialities of reflection. For so long as reflection is bound by the traditions and

prejudices of the hermeneutic situation, it remains hostage to the structures of domination and relations of power which are legitimated *through* these traditions, and which are not transparent to hermeneutic reflection from within them. Critical reflection upon these traditions, for Habermas, cannot appeal to the prejudices of the traditions themselves for the normative basis of its critique. Rather, he insists that it is possible to break out of the linguistic tradition which defines the hermeneutic situation. As evidence, he points to the phenomenon of "systematically distorted communication", from which emancipation can be gained through a process of non-hermeneutic reflection. In the debate with Gadamer, Habermas appeals to the "scenic understanding" achieved in psychoanalysis as a paradigm case of such reflection. In his early work, he calls this kind of reflection "depth hermeneutics".

The scenic understanding which is achieved in the dialogue between the analyst and the patient, it is claimed, retrieves the meaning of an initially incomprehensible "text"; the symptoms of the patient. It does this by appeal to *theoretical* assumptions about psycho-sexual childhood development. By resort to the theory, the analyst can reconstruct an "original" traumatic scene in the patient's early life history, which explains the distorted evolution of the patient's ego-identity. A correct understanding the patient's behaviour is thus conditioned by a knowledge of the causal genesis of the systematic discrepancy between "latent" meaning and "manifest" intention. The crucial point for Habermas is his claim that such knowledge is discontinuous with the language of everyday, traditional horizons of understanding. For in the former case, understanding is guided by theoretical, methodologically non-naive assumptions with explanatory power. The phenomenon of systematically distorted communication is only intelligible if the dialogue situation is not

assumed as always already built into traditions, but is postulated as a normative standard in a *theory of communicative competence*. For Habermas, then, a theory of communicative competence takes over the role of hermeneutic reflection as the *modus operandi* of critique.⁴³

I will examine the conception of crisis which features in this argument in more detail in chapter three. But the significance of its conclusion should be clear; that for Habermas, a theory of communicative competence must be able to identify and to explain the sense in which modernity stands in need of self-reassurance. In (1.1), I sketched the acultural strategy Habermas adopts in reconstructing the logic of learning processes which he takes to characterize modernity. In (1.2), I indicated how crises can be explained in that evolution within the framework of a theory of rationalization. But can the decisive transition out of the *lived* crises of the psychoanalytical patient be explained as a *learning* process within such a framework? In his debate with Habermas, Gadamer doesn't directly address this question. I hope to establish that it is at this point that Taylor's hermeneutic conception of critique advances the debate. From Taylor's perspective, Gadamer can be seen as one of those cultural theorists who inadequately explains rational transitions in the process of self-interpretation. But this will only become evident when the debate on the scope of hermeneutic reflection is recontextualized around the problem of modernity's self-reassurance; or better, when the problem of self-reassurance is reformulated around the principles of hermeneutic reflection.

My aim in this introductory chapter has been to set the stakes for the interpretation and assessment which will follow of the attempts made by Habermas and Taylor to bring the

modern aspiration and struggle for normativity to self-clarification. I now want to explain more precisely how modernity renders *any* claim to normativity problematic. Earlier in this section, I noted Gadamer's objections to the "Enlightenment" view that the methods of the natural sciences hold an exclusive claim to truth. In the previous sections, I emphasized Habermas's point that the resources of human identity, if they are not to degenerate into a "positive" form, must be capable of carrying conviction. If the Enlightenment view is broadly correct, and genuine cognition (of truth) is incapable of carrying sources of human identity with conviction, isn't the project of modernity advanced by the proponents of Enlightenment doomed to positivity? In the next chapter, I try to make sense of the proposition that cognition and human identity stand in irresolvable tension under conditions of modernity.

CHAPTER TWO: THE MODERN TENSION BETWEEN COGNITION AND IDENTITY

(2.0) Introduction

Of contemporary thinkers who insist on the philosophical and normative unassailability of an unqualified project of modernity, perhaps the most uncompromising (and for that reason representative) is Ernest Gellner. The first part of this chapter (2.1) unpacks a remark which is the recurring *motif* of Gellner's philosophical writings; that modernity represents a "wholly new balance between being and knowing".¹ Here, as elsewhere, my focus will be on the coherence of the conceptual resources which are tapped in the articulation of diagnostic claims, rather than on empirical hypotheses concerning the historical or socio-anthropological specificity of what is diagnosed. In my exposition of Gellner, I outline a sketch of the predicaments he takes to follow from a certain way of legitimating beliefs. By adopting what Gellner calls an "ethic of cognition" - within which the *believability* of beliefs is conditioned by the rationality of the procedure legitimating them - moderns bring upon themselves, Gellner proposes, an irreconcilable tension between cognition and identity.

Gellner's understanding of this tension can be provisionally reconstructed around the following conjectures. Human beings have beliefs about themselves and about the world, and the beliefs they have of their place in the world, if true, gives them reason for action. Cognition can be defined as the grasp of truths, and the capacity for cognition reaches fruition when beliefs held in reason are true. If we take it that the capacity for cognition reaches fruition with modern science, what counts as a reason can then be understood in

terms of the procedure which conditions its cognitive success. But when this procedure is applied to beliefs held by human beings about themselves and their *place* in the world - and hence to what could count as a reason for acting - then they no longer seem to admit of truth. The kind of beliefs which are capable of carrying the conviction which cognition affords can only give reason to *instrumental* action in bringing about by the most efficient means a non truth-evaluable end. So in the process of gaining a cognitive status for their beliefs, moderns forfeit conviction-carrying reasons for their own non-instrumental actions. And insofar as beliefs resist assimilation to cognition, they become relativized to the individual (or communal) disposition of the holder(s) of the belief. If, following the terminology of (1.1), we call the resource and context of application of action-guiding beliefs the 'lifeworld', then moderns must confront the predicament of their lifeworld being eroded by cognition. It is a predicament because the beliefs which can be scientifically legitimated are incapable of supporting individual or collective human identities.

After offering a more refined exposition of Gellner's claims, I address the weaknesses in his position by way of reconstructing objections which also serve to introduce the alternatives developed by Habermas and Taylor. In (2.2), I take up what would be Habermas's objections to Gellner's employment of the lifeworld concept, and to his narrowly construed conception of rationality as exclusively instrumental. The fragility of the 'if' which conditions Gellner's identification of rational credibility with scientific legitimacy or instrumental efficacy would be stressed by Taylor as well as by Habermas, though on different grounds. After questioning the significance Gellner attaches to rules in his account of the *intellectual* tension between knowledge and value (2.3), I draw directly on the criticisms which Taylor (amongst others) has put to the anti-realist prejudices to

which Gellner implicitly appeals. My argument at this point is only to draw attention to the ground which would have to be covered for his claims to have adequate support, though I hope, in the course of this, to have put the onus of argument on Gellner's side. The arguments put forward to support the counter-positions of Habermas and Taylor will be rehearsed in later chapters.

What these counter-positions would need to establish by way of reconciling cognition and identity is introduced in (2.4). Gellner acknowledges that the ethic of cognition issues in predicament since it is unable to account for why anyone should abide by it. As has often been noted, in prescinding from value-attribution to the world, the scientific perspective is incapable of attributing value to itself. What I call the Nietzschean strategy for reconciliation embraces this sceptical conclusion, and affirms an ethic of *anti-cognition*. Habermas's and Taylor's strategies can be understood in contrast to this move. Both turn to language in developing a conception of *moral order*, but whereas Habermas attempts to reconstruct it on the basis of the *pragmatics* of linguistically mediated interaction, Taylor seeks to clarify the *ontological* commitments to which the disclosive or expressive dimension of language lends itself. Further, while for Habermas the moral order is essentially open to public recognition, Taylor proposes that under modern conditions, no public realizability of the moral order in which humans are set; in a certain sense, is conceivable. But they share common ground in seeking to counter scepticism by appeal to transcendental arguments concerning a certain kind of unavoidability of *both* cognition and identity. Returning to Gellner's initial insight, both aim to elucidate the problematic balance between being and knowing which characterizes modernity.

(2.1) Gellner on Cognition vs. Identity

Moderns are "doomed", writes Gellner, "to suffer a tension between cognition and identity"². This tension is an inescapable consequence, he thinks, of the consistent and rigorous application of a deeply embedded norm of modern culture; what I shall call 'the validational imperative'. The demand for validation represents the regulative principle of what Gellner calls the "ethic" and the "norm" of cognition.³ According to Gellner, there are two fundamental components to this norm. First, it commands that "anything must be true before it can significantly claim other merits".⁴ The second component requires of truth that it *be* merited in virtue of satisfying maximally risky criteria of epistemic legitimacy. Both components, Gellner claims, issue in a certain "disenchantment" - one which is correctly perceived as threatening to human identity. Faced with this threat, Gellner suggests that philosophers invent a concept of the lifeworld to protect us from the dehumanizing forces of cognition. But this move is self-defeating, he argues, because its motivation is intelligible only under the presupposition of conditions of rationalization incompatible with the applicability of the lifeworld concept. For these are conditions shaped by and fit for instrumental rationality, the very idiom of disenchanting cognition. In this section I will offer an exposition of the grounds for Gellner's thesis that cognition both represents and simultaneously undermines itself as an ethic. In the following two sections I will indicate how the alternative conceptual possibilities for diagnosing the modern tension between cognition and identity offered by Habermas and Taylor can be approached on the basis of weaknesses which can be exposed in Gellner's position.

The first component of Gellner's version of the validational imperative contains the proposition that different kinds of value can be ascribed to beliefs, that one of these is truth,

and that the ethic of cognition prioritizes the truth-value ('truth') of beliefs over other evaluable properties they may have. Besides being evaluable in terms of their truth, beliefs may be held on account of allegiance to a tradition, to a moral or political authority, or to one's inner 'feelings'. Gellner's contrast is between belief systems for which the overriding criterion of acceptability or evaluability of beliefs is their truth, and those for which truth is compromised or overridden by 'idiosyncratic' loyalties to particular traditions, authorities, and faiths. The ethic of cognition requires that, as an ideal, beliefs be held on account of a truth which is independent of the function they serve in perpetuating any tradition other than the growth of objective knowledge. To the degree to which it is possible, the ethic demands that authority and faith ought not "fill out the world", that they should "stand ready to be judged by evidence which is not under their control".⁵ By isolating the truth-evaluability of beliefs, the ethic of cognition impels us to differentiate that function of language which enables us to describe the world felicitously - and hence to issue truths - from functions which, from the point of view of the ethic of cognition, are adventitious to the world-descriptive function of language. It is this differentiation, Gellner thinks, which traditions and the 'common sense' of traditional societies fail to carry through. And it is in virtue of the "systematic conflation of descriptive, evaluative, identificatory, status-conferring and other roles of language", that traditional world-views are "enchanted".⁶ Enchantment, according to Gellner, is a consequence of a vocabulary sufficiently 'thick' to perform each of the different linguistic functions simultaneously; thus giving the appearance of both describing the world and expressing a particular cultural or moral code. The autonomy and primacy of truth appraised by the ethic of cognition, Gellner suggests, serves to alienate "Man the knower" from the "citizen and the moral being".⁷

In order to maximize the extent to which beliefs are held in virtue of their truth, and to minimize the degree to which non-assertoric functions of language serve to influence the content of belief, the ethic of cognition requires of truth-claims that they pass a maximally rigorous selection procedure. It *may* be the case that traditions and common sense are the vehicles of truth-content, but the ethic of cognition biases the burden of proof by invoking criteria of legitimacy which are minimally dependent on particular cultural contents. The second aspect of Gellner's validational imperative does not reduce truth to legitimacy, rather it insists that since it is *better* to hold beliefs which are true rather than false - and this independently of the consequences of holding them - we are *obliged* to maximize our chances of arriving at the truth, and this means following a procedure for reaching them which minimizes the risk of error. We are obliged, that is, to be able to legitimate beliefs in as stringent a manner as possible. The task of epistemology is to clarify how this obligation is best met.

Gellner proposes that once the task of epistemology is seen in this way, the preoccupation with the 'foundations of knowledge' in the Descartes-Kant canon of modern philosophy appears not so much as an attempted explanation of cognitive success and its possibility, but more as a concern for outlining a programme of "recommendations for the proper conduct of our intellectual life".⁸ The significance of epistemology, then, lies in the series of "cultural injunctions" it elaborates, the most important being the validational imperative to place all beliefs and practices *sub-judice*. For instance, Locke's description of the role of philosophy as the "handmaiden to science" can be read as motivated not just by the perceived need for science to be given foundations, but by the understanding that the emerging science of his time faithfully satisfied the requirements of proper intellectual

conduct; conduct which its philosophical handmaiden could assist.⁹ But modern epistemology, at least insofar as it is set on its way by Descartes and Locke, could not explain the cognitive advance of modern science, Gellner suggests, because by focussing on the *individual's* acquisition and justification of beliefs, it failed to grasp the broader cultural injunctions which came to inform the emerging scientific *system* of belief. So Gellner holds that the terms of acceptability of beliefs can systematically differ, that with modernity a cultural injunction emerged that beliefs be acceptable - that is, worth having - if one can generally expect them to have satisfied certain criteria of legitimacy, and that these broader cultural criteria can serve both to demarcate scientific thought, and to account for its cognitive advance.

What, then, are these criteria? Gellner's claim is that modern cognitive practices are peculiarly constrained in the provision of explanations which are open to public and repeatable testing. This formal norm of genuine cognition, according to Gellner, serves to neutralize the cultural "cocoon" within which, in cognitively "low-powered" societies, knowledge claims are protected.¹⁰ It is formal in virtue of being maximally, if not absolutely, culture-unspecific. Gellner can call this culture-neutral baseline of explanation "mechanism" because "a machine is an artifact which can be reproduced at any time, in any place, in any society, provided that the same specified materials are used and put together in a publicly specifiable way".¹¹ The ethic of cognition requires the same of legitimate explanations. But in order to satisfy this requirement, it must avail itself of certain conceptual resources. Most significantly, these resources must be strictly rule-bound. The concept of a rule, Gellner suggests, captures the deep significance of the norms of public specifiability and repeatability. Rules command consistency and symmetry, "like

cases are treated alike" in accordance with them - they do not admit of idiosyncratic variation. The disposition towards mechanism which defines the scientific mode of cognition can thus be interpreted as a rigorously sustained bias toward rule-boundedness. Procedures of explanation must follow an orderly, rule-governed method, and the concepts by which they are articulated must be of such a kind as to be applicable in a strictly ordered, non-idiosyncratic way.

But a presupposition of treating like cases alike in accordance with rules is that different kinds or classes of cases be distinguished and differentiated from each other. In the case of cognition, the effect of such differentiation is to purify the language of possible explanation from all other compromising linguistic functions. The subsequent disenchantment is reinforced by what Gellner calls an "ethic of rules", constraining the knower to a vocabulary of a kind which is expunged of human idiosyncrasy. The order, regularity, and symmetry imposed by rules on the behaviour of objects and concepts by mechanism, disenchant them by ruling out the spontaneity and idiosyncrasy characteristic of agency. The world loses the meaning discharged by the kind of concepts through which human agents must understand themselves as agents. Gellner observes that the loss of spontaneity and freedom which Weber captured in his image of modern, bureaucratically organized society as an "iron cage", can thus be extended to the modern scientific conceptual organization of the world.¹² A language of genuine cognition is gained at the expense of a vocabulary in which the identity of free and meaningful human agency can be recognized and expressed: the concepts "in terms of which identities are forged and life is lived" are submitted to an irreversible form of "intellectual erosion".¹³ To adopt a formulation from John McDowell, modern science develops conceptual capacities which are

directed to the kind of intelligibility that is proper to the realm of *law*, precisely by separating that intelligibility from the kind that is proper to *meaning*.¹⁴ But since human beings do fall within the realm of law, Gellner sees no alternative to the view that "whatever is worth saving in our conception of ourselves needs to be reconstructed in terms of conceptual apparatus suitable for characterizing the realm of law as such".¹⁵ This is just the constraint required by the mechanistic baseline of legitimate explanation, which thus issues, Gellner believes, in the modern "certainty of reductionism"; "that everything is an unedifying something else."¹⁶

So the modern tension between cognition and identity, as Gellner presents it, results from a conflict between the validational imperative and the human need for enchantment. One way of dealing with this conflict, which Gellner believes to have been taken by many contemporary philosophers, is to re-establish identity and self-reassurance by separating off or *bracketing* the world in which humans live and endorsing it, as it were, from within. It is in the service of re-enchantment that Gellner sees the significance of the notion of the lifeworld. The lifeworld (*Lebenswelt*) - "the ordinary world in which we conduct our daily life"¹⁷ - provides both the resources for identity-formation, and the field of expression of human identities. It becomes problematic, he contends, with the differentiation of the languages of cognition and life, with the discontinuity between the vocabularies of science and the everyday. This is partly because competence in the use of the former generally requires a specialized and technical training, but the point Gellner emphasizes about the modern lifeworld is that it is partly *constituted* by this discrepancy between the language of genuine cognition and 'ordinary' language. That is, the modern recognizes the "interim status" of the claims of common sense as objects of possible "re-validation" by science, and

thus of possible reconceptualization according to apparatus suitable for rendering them intelligible within the realm of law. This lack of congruence is evidenced, he suggests, by the modern preference for the "idiom for which we have greater cognitive respect" - namely, science - when dealing with problems of grave practical consequence.¹⁸

The paradox of the preference for scientific problem-solving, assuming for the moment that it is one which generally obtains, is this; that although the idiom of lifeworld thought lacks competence for dealing with issues of real gravity, it is also responsible for *determining* what these issues are to be. The implication of Gellner's position is that within the modern lifeworld, what is to count as of real importance is a matter of something external to it - namely, the sphere of expert scientific/technological knowledge and practice. It is part of the identity of the modern lifeworld that issues of greatest importance be specified in a way amenable to scientific or technological resolution. But science and technology deal with a world devoid of meaning, or at least must presuppose such a world in order to reap the goods demanded by the modern lifeworld. Consequently, Gellner argues, under the pressure of the validational imperative, the boundary around any putative lifeworld collapses, such that the very invocation of a lifeworld testifies to its non-availability as a real option.¹⁹ The separability of the lifeworld from the world of cognition, and of the radically different idioms in which they are articulated, Gellner asserts, is enough to undermine the ultimate legitimacy of the claims of the lifeworld. Yet it only makes sense to impute a lifeworld under modern conditions, as a measure to preserve human identity from the anonymity of the world-machine. Speaking on behalf of the modern, Gellner claims that the lifeworld can only properly be imputed with irony, since "we have become aware of it when we no longer live in it, at any rate not exclusively or predominantly".²⁰

The lifeworld becomes 'eroded' by an idiom of cognition 'detached' and 'autonomous' from it; the latter destabilizes the former irreversibly.

Gellner is not referring here simply to an 'intellectual' erosion, but to one which he thinks has its roots in the increasing rationalization of action in the modern world. Gellner's account of how it occurs can be summarized more clearly if we keep distinct what Gellner himself conflates - the two senses of lifeworld distinguished above. The lifeworld is both the *source* of identity-formations, and therefore of categories and judgements about what is most worthwhile and important. It is also the means of application and *reproduction* of these values in everyday life. Keeping these two senses separate, Gellner's account of how the lifeworld becomes eroded goes as follows. The practical application of cognition (technology) has a disintegrative effect on the lifeworld (in the second sense) because of the idiom of its operation. Technology aims at the most efficient means of solving a pre-determined, closed problem. Likewise, the science which is applied operates by means of hypotheses which seek to explain a given problematic state of affairs. The idiom of cognition, then, is applied in everyday life in order to solve problems by way of discovering the most efficient means of achieving a given, pre-determined end. It requires the maximum degree, for Gellner, of instrumental rationality.

The implications of this for the lifeworld (in the first sense) become apparent when Gellner discusses the possibility of judging the rationality of the conduct of a life as a whole. Against the background of the modern worldview, nature is an empty resource for judgements concerning the ends of human action - the world is a morally neutral mechanism. But it does behave in an orderly, rule-governed way, which affords a basis for

judging the rationality of actions in terms of their efficiency in being instrumental to the realization of given desired ends. Particular ends of action can also be judged rational to the degree that they are mutually consistent, so that the efficient realization of one specific end might be judged irrational if it is incompatible with the realization of another, more desired end. Even here, of course, the rationality of the end is being judged as a means to a further prioritized end; that is, in terms of a relative instrumentality. But what about the rationality of the end for the conduct of a life as a whole, where the question of relative instrumentality does not arise (we only live *one* life as a whole)? The problem here, Gellner observes, is that the presuppositions of instrumental rationality break down, for two related reasons. First, the specificity required for an end of action to be the basis of a calculation of efficient means, renders any such end unfit for the purpose of a whole life. By way of illustration, Gellner remarks that where the desired end is a holiday partner, the qualities sought in that person are readily specifiable, and a decision can be reached which will realize the desired end with instrumental reason. But in the case of life-long commitments, such as the choice of a spouse, no such qualities can be readily specified. Second, and more crucially, the diversity and plurality of plausible fundamental ends is inconsistent with the *unitary* identity which conditions instrumentally rational choices. In the following way, Gellner proposes that this problem assumes a much broader significance under conditions of modernity.

Gellner coins the phrases "Fixed and Variable Cognitive Capital" to illustrate the changing significance of instrumental rationality in the modernization process.²¹ The expression 'Cognitive Capital' serves as a reminder that bodies of knowledge are available to us which are a useful resource for dealing with particular life problems. It is 'Fixed' if the system

of beliefs which make it up is relatively rigid and stable, and so immune from the challenges of recalcitrant experience. Accordingly, it is important that these beliefs be *sui generis*, if they are to be understood as hypotheses at all. Further, it is "the concepts and ideas of this Fixed Cognitive Capital" which serve to articulate and legitimate "identities" and "personal relationships", as well as power and hierarchy structures. It is thus the *resource* for the legitimation of personal and social identities, and also a conservative counter-weight to change. Where there is a greater degree of evidence sensitivity, problems are rationally resolvable in terms of the "Variable Cognitive Capital". This is the domain of instrumental rationality, where hypotheses are constructed to discover the most efficient means to a particular end. In bringing everything under the hypothesis, scientific thought destabilizes the "Fixed Cognitive Capital" out of which the value of instrumental rationality arose:

The modern scientific/industrial world is simultaneously driving us in two incompatible directions. By eroding the old frameworks and requiring neutral, homogeneous legitimation of beliefs, it pushes the world into becoming a Bundle of Hypotheses, and thus a home fit for instrumental rationality... At the same time, this extension of the Bundle of Hypotheses and the corresponding reduction of the rigid framework also eventually make rational calculation harder in many areas where it is now expected and which were previously exempt from it. The more general or fundamental features of the world, though now demoted to the status of mere hypotheses, often elude rational assessment because they are unique or *sui generis* or very fundamental. They become relativised, optional, and deprived of their privileged, entrenched status - but without becoming, for all that, eligible for rational, instrumental evaluation.²²

This, I think, is the closest Gellner comes to diagnosing the real source of the modern tension between cognition and identity. I shall now consider reasons for thinking that it stands on an inadequate conceptual foundation, by way of reconstructing the kind of response which Habermas and Taylor would make to it.

(2.2) Morality and Cognition - Habermas's Response

One way of responding to Gellner's position would be to retain the validational imperative, while challenging Gellner's formulation of it. One might want to accept that the cognitive achievements of modernity presuppose a systematic differentiation of descriptive, prescriptive, and evaluative linguistic functions; that the terms of acceptability of belief take on an increasingly formal and reflexive character as legitimation is secured with minimal appeal to pre-reflective, tradition-specific dogma; and that there is a sense in which this process renders the claims of the lifeworld problematic. But if the validation imperative is dissociated from its positivistic construal as mechanism - if it is *radicalized* to cover the claims of the lifeworld themselves - then the imputation of the lifeworld can in turn be divorced from the motivation for re-enchantment. If the first component of Gellner's version of the validational imperative is dropped, and cognitive worth is attributed to redeemable validity claims *other* than truth-claims, then a corresponding shift in our conception of the idiom of validation (the second component) is required. Then the effect on the lifeworld of different kinds of rationality would also need to be distinguished, and the tension between cognition and identity diagnosed by Gellner retheorized. It is by such a radicalization of the validational imperative that Habermas can be seen to respond to Gellner's position. The motivation for this move can be reconstructed from the need to overcome weaknesses primarily in (1) Gellner's employment of the lifeworld concept, and (2) in his understanding of instrumental rationality.

(1) There are several reasons for thinking that Gellner's polemic against the concept of the lifeworld is unsatisfactory. First, Gellner equivocates in his employment of the notion; he oscillates between the mutually incompatible views that the claims of the lifeworld have no

cognitive status at all, and that this status is cognitive but precariously so. The first view amounts to an a priori rejection of the cognitive capacity of most of natural language, and would be in need of much further metaphysical support. Not only is such support not forthcoming, but it is difficult to see how it could be given the radical epistemological constraint on metaphysical thought built into the validational imperative. But if, for epistemological reasons, the claims of the lifeworld are ascribed a precarious cognitive status, one will want to know more about why the validational imperative *itself* is excluded from them. To appreciate this point, it is enough to take a brief look at how the concept of the lifeworld gets its philosophical significance.

As Husserl presents it, the lifeworld is the background or "horizon" of unthematized cultural certainties which is presupposed by, and always pre-given to, the human enquirer. The lifeworld is "always already there, existing in advance of us", presenting objects to "always somehow practically interested subjects".²³ Because of this entwinement of theory and practice within a lifeworld, even the "knowledge of the objective-scientific world is 'grounded' in the self-evidence of the life-world"²⁴. This self-evidence is irreducibly intersubjective, its taken-for-grantedness is constitutively shared by members of the living community of which the theoretician or scientist is always a part.²⁵ The lifeworld is a 'world' which "is always prior to the subject that relates itself to objects in knowing and acting"²⁶, it provides that background context against which a subject can confront an objective world in an appropriate epistemic attitude. Gadamer's hermeneutics, Habermas observes, transformed the lifeworld concept into one of "a culturally transmitted and linguistically organized repository of meaning patterns"²⁷ which is pre-given as a linguistic and cultural horizon for theoretically and practically engaged subjects.

The lifeworld considered as an epistemic totality corresponds closely to what Wittgenstein calls "the sub-stratum of all my inquiry and asserting"²⁸, and "the inherited background against which I distinguish between true and false".²⁹ It is therefore not something which *as a whole* I can stand outside of and judge *as* true or false, and thus not something which is an apt object of doubt. Certainty is properly conceived as a function not of 'intrinsically credible' beliefs which can be construed as the foundations of knowledge, but of a mutually reinforcing, holistically structured web of thought and action. It is against this certain - but tacitly-known - background that particular doubts and knowledge claims are tested and validated. And it follows from this that not all propositions can be hypotheses since the very possibility of judgement presupposes commonly accepted standards or *norms* of judgement.

Such considerations play an important part in the use Habermas makes of the category of the lifeworld. The lifeworld is a horizon of "more or less diffuse, always unproblematic, background convictions".³⁰ Beliefs, assumptions, definitions and expectations present themselves in an unthematized, intuitive, pre-given way in the lifeworld horizon. Problematic beliefs and situations are "*encompassed* within the horizons of a lifeworld"³¹, but the lifeworld as such "cannot become problematic, it can at most fall apart". The lifeworld as a whole is therefore immune from total revision. It is encountered by subjects as a pre-given, pre-interpreted reality the limits of which "cannot be transcended"; for the lifeworld is that "transcendental site where speaker and hearer meet".³² In employing the lifeworld concept in this way, Habermas is exploiting a function which the notion has always served; to signal the exhaustion of foundationalist epistemology. From Husserl to Habermas, the idea of the lifeworld is a philosophical tool constructed for the purpose of *saving* cognition from the *reductio* which issues from sceptical demands for foundational

justification. In other words, it serves to redeem the validation imperative from the incoherence into which it collapses when pushed beyond its proper scope of application. Gellner himself acknowledges that the unbounded application of the principle of validation threatens to undermine all cognitive claims by infinite regress, but recommends that we resign ourselves to this predicament.³³ But by rejecting the lifeworld concept *tout court*, he allows himself no resource for making a case for such an acknowledgement. And this refusal is a consequence of a misunderstanding of the philosophical motivation behind the concept.

Contrary to Gellner's presumption, the lifeworld need not be a safe haven for philosophers on the run from the world-machine. This is particularly clear in the use Habermas makes of the concept. Habermas is careful to distinguish world concepts which are the referential presuppositions of redeemable validity claims from the lifeworld concept which represents the context which conditions the meaning of those claims. According to Habermas's theory, I can raise claims with presuppositional reference to an objective world ("the totality of objects and states of affairs"), what he calls a social world ("the totality of legitimately regulated interpersonal relations"), and a subjective world ("the totality of experience to which a speaker has privileged access and which he can express before a public").³⁴ These 'worlds' are the ontological correlates of claims which can be validated by appeal to the irreducible criteria of truth, rightness, and sincerity or truthfulness respectively. The lifeworld, on the other hand, plays no such ontological role; speakers and hearers cannot refer to it in the way they can to the objective, social, and subjective worlds. Hence knowledge of the objective world does not compete with the lifeworld as it does in the scenario depicted by Gellner.

A third objection can be directed towards inadequacies in Gellner's diagnosis of the tension between cognition and identity arising from his abandonment of the lifeworld concept. In order to distance himself from projects of re-enchantment, Gellner chooses to articulate the threat posed by the idiom of cognition to the linguistic resources required for making sense of and expressing human identity as disintegrative of the lifeworld. The concept relinquished, he is forced into coining terms like "Fixed and Variable Cognitive Capital" for the purpose of explaining this effect. But Habermas's lifeworld concept can do the same job, only much more elegantly and with greater explanatory power. Besides being a resource of cultural or epistemic certainties - the pre-given stocks of knowledge upon which speakers and hearers draw - there are also institutional orders which regulate group memberships, as well as individual pre-theoretical skills and competences constituting the horizon of the lifeworld. So for Habermas, the lifeworld is the source not only for the production and reproduction of knowledge, but also of social solidarities and individual personal identities. Under the pressure imposed by the validational imperative, the lifeworld itself can and does become rationalized. With the concept of a rationalized lifeworld, Habermas can capture both the thought that traditionally sanctioned stocks of knowledge are put at risk under the requirement for validation, and that any particular validation is encompassed by a background totality of taken-for-grantedness which cannot be bracketed at will. Not only does this move put Gellner's charge of 'bracketing' on the other foot, but it makes room for a conception of rationalization which covers each of the three dimensions of the lifeworld.³⁵ And as we shall now see, this sheds a different explanatory and diagnostic light on the significance of instrumental rationality.

(2) Habermas would argue that from the point of view adopted by Gellner, the notion of a rationalized lifeworld falls on a blindspot for which a narrowly constricted conception of rationality is responsible. The notion of a rationalized lifeworld, then, must be complementary to a conception of action which is non-instrumentally rationalizable. If Gellner were to have recourse to a model of action the rationality of which is not determined by instrumental success - by the criterion of efficiency of means to a non-rationally decidable end - he could avoid what would then appear as a metonymic fallacy of construing one particular moment of rationality as the whole. With what Habermas calls communicative action, he claims to have just such recourse.

A lifeworld becomes rationalized, in Habermas's sense, "to the extent that it permits interactions which are not guided by normatively *ascribed* agreement but - directly or indirectly - by communicatively *achieved* understanding".³⁶ A communicatively achieved understanding is one which is reached purely on the basis of the better argument, whereas an agreement is normatively ascribed if it is accepted habitually or uncritically - say, on the basis of some unchallenged convention, authority, or tradition. Communicative action is furnished by and also reproduces the background horizon of the lifeworld, while putting at risk the particular claim which is either implicitly or explicitly raised in the action. The understanding to which communicative actors are oriented is one which is *meritorious* of rationally motivated intersubjective recognition. To say that a lifeworld can possess degrees of rationalization, is to propose a thesis concerning the scope made available for communicative action.

I will offer a more detailed discussion of Habermas's concept of communicative action in

chapter four. For the moment, what matters is the sense in which the conception of rationalization for which communicative action is the vehicle differs from and marks an improvement upon Gellner's thesis concerning the erosion of the lifeworld by instrumental rationality. While Habermas's claim converges with Gellner's view that the validational imperative puts the claims of the lifeworld at risk, it diverges in expanding the scope for validation beyond the constraints imposed by 'mechanism'. The privilege which Gellner ascribes to scientific method (the ethic of cognition) corresponds to a foreshortening of the rational potential of communicative action. Communicative reason finds its criteria "in the argumentative procedures" for redeeming "validity claims geared to intersubjective recognition".³⁷ But Gellner's exposition of these procedures is too narrow. Not only does it fail to account for the possibility of criticizable - but non-truth-evaluable - validity claims, but it does so by undercutting the role of everyday, uncoerced dialogue as the idiom of validation. Once the latter is taken as paradigmatic, then the human capacity for "making true statements and implementing plans"³⁸ loses its privilege, and the space emerges for replacing an ethic of cognition with a more fundamental communicative ethic, oriented by the norm of mutual recognition. Further, it enables Habermas to distinguish two different senses of rationalization which Gellner assimilates, leading to a misleading account of the source of tension between cognition and identity.

The lifeworld concept gets its distinctive meaning in Habermas's theory not only from its complementary relationship to communicative action, but also in its distinction from social reality considered as a 'system'. As we saw in (1.2), society *qua* system is rational to the degree that it integrates action consequences according to criteria of efficiency in performing functions necessary for its self-preservation. We saw that in the subsystems of

the economy and the bureaucratic state, action integration is mediated not communicatively, but by the delinguistified steering media of money and power, which react back and mediatize, instrumentalize, and colonize the symbolic reproduction of the lifeworld. Accordingly, the main cause of lifeworld erosion is not so much disenchanting cognition, but the pseudo-communicative or delinguistified media of system integration. Habermas acknowledges that cognitive claims become increasingly rarefied and split off from everyday discourse in the domain of scientific expert spheres, but the specialized training which is required to understand these claims *also* conditions competence for participation at the highest level of discourse in which other validity claims are thematized; moral and aesthetic discourse also get separated from the horizon in which everyday life is led. But this impoverishment of identity is of secondary significance to the colonization of identity wrought when the horizon within which identities are forged become mediated by money and power. It is the latter which "deworlds" the lifeworld, not cognition. But a symptom of colonization is a one-sided rationalization of the lifeworld in its cognitive/technological dimension, and herein lies the *imbalance* between being and knowing characteristic of the times.

In short, Habermas would argue that Gellner's fundamental error can be traced back to the central flaw in Weber's "disenchantment thesis"; a false opposition between the identity-consolidating Reason built into religion and metaphysical world views, and an identity-resourceless instrumental rationality built into modern forms of action and scientific knowledge.³⁹ He replies by way of reconstructing the validational imperative in a manner which drops the priority it cedes to truth, and which substitutes procedures of argumentation operative in everyday communication for the mechanistic idiom of validation. It follows

that Habermas's idea of a rationalized lifeworld stands or falls with the availability of a correspondingly expanded conception of rationality, the criterion of which is not identified with the kind of instrumental success which betokens the increasing rationalization of system-integrated action. Further, if what is "de-worlding" about money and power is that they are "delinguistified", Habermas will have to extract from language a model of normativity which explains the vulnerability of human identity to this kind of corrosion. He needs to establish not only that the claims of morality have as good a place in argumentative procedures as scientific/cognitive claims, but also that there is good reason for abiding by these procedures - that doing so satisfies the requirements of human identity.

(2.3) Morality and Identity - Taylor's Response

A different way of challenging Gellner's position would be to retain the truth-requirement of beliefs relevant for sustaining human identity, but to divorce it from *sub-judice* procedures of validation. Rather than radicalizing the validation imperative to cover obligatory norms of action, this move incorporates cognition into identity and thereby *ontologizes* it. The validation imperative is taken as one form of human self-interpretation amongst others; its status as a cultural injunction is affirmed, but it is *also* allowed to admit of truth, conceived as the felicitous disclosure of a moral world. One might want to reject the view that the systematic differentiation of descriptive and evaluative linguistic functions necessarily does constitute cognitive advance, as well as the view that beliefs merit rational acceptability to the degree to which they transcend their culture-specific content. The objection here would not be - *pace* Habermas - that the privilege Gellner accords to the assertoric mode of language is arbitrary from the perspective of the validation imperative (properly conceived), but that the domain of assertion is arbitrarily restricted from the

perspective of what might be called the 'interpretative imperative'. The restriction on interpretation imposed by mechanism can appear as arbitrary in the light of further objections which can be put to (1) the significance of rules, (2) the mechanistic constraint on the language of explanation, and (3) the concept of rationality, as they feature in Gellner's account.

(1) Gellner correctly relates rule-boundedness to consistency, publicity and repeatability. He then takes these as criteria for demarcating scientific cognition from magic, on the grounds that the latter allows of explanations which are inconsistent with each other, and which are protected from public and repeatable testing procedures. Gellner takes this consideration to justify the claim that our trust in scientific knowledge is of a fundamentally different *kind* to that which not only magic, but also the knowledge claims implicit in the discourse through which sense is made of the living of lives, enjoins. But is this latter claim justified? Can the concept of a rule do the work Gellner requires of it? What first needs to be observed is that all discourse qua discourse is rule-bound, and not just in virtue of syntactic structure (which has no bearing on Gellner's claim). Discourse is also rule-bound insofar as the concepts which articulate the propositional content of a sentence are capable of being applied correctly or incorrectly. Now if this is a property which all concepts which can feature in meaningful sentences have, then it is absurd to say that the conceptual resources available for employment in legitimate discourse can be 'eroded' by rule-governed constraint. This is the case even if by 'legitimate discourse' we mean 'fact-stating' discourse, since any grammatically well-formed indicative sentence has *prima facie* propositional status - having truth conditions which when satisfied would make it true.

Does Gellner give us any reason for overturning the prima facie propositional (truth-evaluable) status of grammatically well-formed indicative sentences, *whatever* their conceptual content? It is not enough, as Gellner suggests, to say that the rule-governed peculiarity of natural scientific discourse lies in the fact that 'like cases are treated alike', since what counts as a 'like' case is itself what a rule needs to determine. His position must be either that what counts as the 'same' instance of the correct application of a concept is fixed in advance of and *independently* of our practices of explanation, or that rule-boundedness is itself a function or expression of those practices. The former position would commit him to the claim that there are instances of the 'same' facts which are graspable in a linguistically unmediated fashion but this view is inconsistent with his insistence that "it is our practices of explanation which disenchant us".⁴⁰ However, it is a view which the analogy he draws between a culture-neutral baseline of explanation and the operation of a machine might tempt him to make. As Gellner described it, a machine is something which is reproducible "at any time, in any place, in any society, provided that the same specified materials are put together in a publicly specifiable way". Not only does drawing this analogy with legitimate explanations beg the question of what counts as "the same", but one might further object to the decontextualization proposed in the analogy. Gellner's picture of a machine, the Wittgensteinian would reply, misleads us "into conceiving justification as the unfolding of the pre-determined properties of a chain",⁴¹ and this because the supposed working of the machine "is only the *picture* of the working of a machine".⁴² Continuing the analogy, if all we have is a picture of a machine, we are left with the crippling sceptical paradox of how a rule can be *followed* determinately, since any way of 'going on' from a rule - i.e. in the 'same' way - specified independently of its applications, can be made to fit the rule.⁴³ But what we actually have are not pictures, but *practices* of

rule-following, and these practices are always context-bound. Gellner's machine-imagery thus lends a false impression by imputing a baseline of explanation which is divorced from contextualized practices of explanation.

If, on the other hand, Gellner holds that rule-boundedness is an expression of our practices of natural science, then one will want to know why other discourses do not make the grade for fact-stating. His argument now seems to be that this is so in virtue of the degree of rigour in rule-following required by legitimate participation in the scientific language game. But this argument only establishes that discourses exhibit differences along a continuum of rule-boundedness, a continuum which covers various degrees of consensus regarding what counts as rule-bounded, and what counts as a fact. The relative *flexibility* of the criteria which govern the correct application of evaluative concepts - and to that extent their idiosyncrasy - also serves to confirm their resistance to arbitrary usage. But if we take this *resistance* to be the tell-tale sign of fact-stating in scientific discourse, then we have been given no reason for thinking that the same does not hold for evaluative discourse.

But perhaps these objections miss Gellner's point. He might reply that the symmetry and lack of idiosyncrasy which is required of the behaviour of concepts which feature in genuine cognition is elucidated not so much by their rule-boundedness, nor by the public and repeatable verification procedures of the theories containing them, but rather by something akin to the idea of *primary qualities*. Primary qualities are those properties or powers the existence of which is not wholly contingent upon the existence of beings (like humans) who are disposed to be affected by them in the form of experience. They are distinguished from secondary properties which exist solely in virtue of such dispositions. If the world

contained no beings with the idiosyncratic sensory (or linguistic) constitution of humans, it would contain no secondary qualities; whereas the existence of primary qualities is quite independent of such human idiosyncrasy. The crux of the distinction, however, comes from the further supposition that human nature is prone to the error of *projecting* what is relative to the idiosyncrasy of the human enquirer onto the world as it exists - so to speak - absolutely. Hence there issues an epistemological and indeed ethical requirement to disengage from these human idiosyncrasies; to gain a perspective on the world from a vantage point which is neutral with respect to the peculiar significance it has for humans; to give an absolute account of reality on which there would be (in principle) universal agreement between all enquirers who had successfully managed to overcome their naive, partial, idiosyncratic standpoints. Such an account would be written in a language of primary quality concepts, and only explanations couched in them would qualify as genuinely cognitive.⁴⁴ Gellner's ethic of cognition, and his version of the validation imperative, could then be translated as: "*Don't project!*". But this move raises its own difficulties.

(2) One might start by challenging the coherence of the primary/secondary quality distinction, and of the very idea of an absolute account of reality. But this is not the path which Taylor takes.⁴⁵ Rather, he argues that while the absolute conception *does* have a proper scope of applicability, insuperable difficulties face it when *overextended*. Taylor would argue that the shortcomings in Gellner's diagnosis of the tension between cognition and identity result from just such overextension.

One difficulty arises as soon as we put the question; are there forms of explanation which are not adequately articulated or conceptualized in absolute terms? One obvious case would

be the explanations we give of human behaviour in terms of the moral significance attachable to what motivates the behaviour. Human beings apparently lead their lives in a world which has moral significance for them, and we commonly understand and explain how these lives are lived by appeal to such notions as "honestly", "courageously", "sensitively", "magnanimously", "spitefully", "inauthentically", and the like; what Taylor calls "desirability-characterizations" (or in the latter two cases, "undesirability-characterizations"). Such notions have no place in the absolute conception of reality, which accounts for the way the world is independently of the meaning it has for humans.

Now if Gellner's thesis is correct that the terms invoked in the explanations through which humans lead their lives require revalidation by science, then one will want to know how the absolute conception can cover and *improve upon* this *prima facie* explanatory language. What seems to be required is a translation of these terms into a vocabulary which separates the (neutral, human independent) reality of the situation described, from the experience of that reality (say, in terms of an intrinsically neutral - since applicable to a neutrally describable fact - 'pro-' or 'con-' attitude), and which when combined covers in scientifically valid terms the naively ascribed quality. But objections of various strength can be put to this requirement. First; as a matter of fact, no such vocabulary is available, and it is difficult to imagine what it would sound like. Second, it fails to give adequate account of the phenomenology of moral experience, which seems to inform us that there can be fit or apt objects of this experience, and hence that the 'fact' which it is construed as projecting upon is *non-contingently* related to it (and hence that the fact is not neutrally or 'absolutely' describable); to describe the situation which is the object of the experience (or concept) in absolute terms is to change the meaning of that experience (or concept), since there seems

to be an essential link between the facts which make up the situation and the attitudinal response to it which is captured in the original term.⁴⁶ The meaning of the term does not tolerate breaking the link, since the putatively "revalidating" description is not co-extensional with it. One might conclude from this that we need not accept Gellner's thesis that modern science or genuine cognition submits the concepts in terms of which identities are forged and life is lived to an intellectual erosion, but rather that the imputed erosion is a *pseudo-intellectual* one. If we take it upon ourselves that our intellectual life is best conducted under the recommendation "don't project!", we can be made to *think* we are projecting reactions onto neutrally describable facts, even when no such description is available, convincing, or even conceivable. In other words, the injunction gives a distorted, or at least inadequate, model of proper intellectual conduct.⁴⁷

This second objection can be put another way. Modern scientific explanations, guided by the claim to an absolute account of reality, seek to improve upon naive everyday explanations by identifying facts independently of idiosyncratically human reactions to them. But by taking this neutral stance, by disengaging from the reactions which typically 'accompany' the understanding of a phenomenon, it is possible that we can lose our grasp of the phenomenon which needs explaining.⁴⁸ In the case of moral phenomena, the assumption of the neutral stance can *disable* us from arguing competently about them. This suggests that the method of prescinding from moral 'prejudices' - or the horizon of meaning within which we are engaged in the leading of our lives - does not necessarily give better insight into the domain of reality under consideration. So the absolute description, rather than giving us the really true story (Gellner's "certainty of reductionism"), actually loses from view some aspect of reality requiring explanation. If this is the case, it seems that it

is *irrational* to predicate descriptions of this domain couched in absolute terms as cognitive.

(3) But in order to be able to say this, we need a richer concept of rationality than Gellner provides. Taylor suggests that the concept of rationality is best captured by the idea of perspicuous articulation. A perspicuous articulation will be well-ordered and consistent; it will give an account which clearly distinguishes one kind of case or phenomenon from another, in formulations which are consistent internally and with each other. I have a rational grasp of something if I can articulate it in a perspicuously ordered, and a fortiori consistent account. Likewise, an action merits the ascription rational if it is consistent with achieving some more or less clearly defined objective which gives it its point, and it is irrational if it frustrates the achievement of that goal. But the criterion of consistency does not exhaust the concept of rationality; consistency is a necessary but not a sufficient condition of a rational account, or a rational action. For the canons within which an account is consistent, and the objectives which are consistently met by means of action - it at least makes *prima facie* sense to say - can also be more or less rational.⁴⁹ For canons and objectives often appear as rivals, and in virtue of this, assessable against each other by appeal to some standard other than the formal one of consistency. Taylor allows that where the subject-matter admits of it, the demand of rationality is to adopt a disengaged, neutral perspective. But where the disengaged perspective issues in *inarticulacy*, as is proposed in the previous objection, it is *to that extent* irrational. The rationally superior account would be the one which renders articulate what was otherwise confused or occluded from view. Such an approach to the concept of rationality has two crucial implications. First, it leads us to think of standards of rationality in terms of *contrasting* degrees of perspicuity possessed by available vocabularies, rather than in terms of some neutrally describable ideal

standard. Second, it encourages a conception of truth as what is *disclosed* by a perspicuous articulation, rather than in terms of correspondence or consent.

From Taylor's standpoint, the most glaring inarticulacy in Gellner's account is its failure to render perspicuous the nature of the ethic of cognition as an ethic. It espouses the desirability-characterization of adopting a disengaged, neutral stance towards the world - that it is *better* to hold beliefs from that perspective and in that attitude - while neglecting to account both for what is involved in attributing such characterizations, and for the possibility of attributing them *correctly*. In the terminology developed by Taylor, it is just such characterizations which constitute a human identity. Gellner's ethic of cognition is thus best understood as an expression of identity, one deeply embedded in modern culture, but which undercuts itself when applied to itself as an identity - since the concept of an identity is not available from the disengaged perspective. Accordingly, the tension between cognition and identity diagnosed by Gellner should rather be attributed to something *internal* to the modern identity itself.

(2.4) Strategies for Reconciliation

In *Legitimation Crisis*, Habermas takes up the predicament generated by the modern tension between cognition and identity with the following question; "If world-views", he asks, "have foundered on the separation of cognitive from socially integrative components", if such world-views "today belong irretrievably to the past", then what else can fulfil "the moral-practical task of constituting ego- and group-identity?".⁵⁰ Is there scope, he asks, for a morality without roots in "cognitive interpretations of nature" which could "adequately stabilize itself" and "secure the identities of individuals and collectives"? If we assume that

the cognitive/scientific task of finding truths about the world is not expected to deliver anything of moral significance, and that moral significance is something which is presupposed in the identity of human beings, do human beings under such conditions really have a sustainable identity at all? If the socially integrative fabric of a moral order no longer belongs to the natural order, does it make sense to talk of *any* order by legitimate appeal to which human beings can satisfy their moral-practical needs and tasks? The objections put to Gellner's position embryonically express two different strategies for tackling this question, and hence for reconciling modern scientific cognition with human identity. I will now offer an outline of the strategies adopted by Habermas and Taylor, but I shall begin by introducing a third against which both can be contrasted; that proposed by Nietzsche.⁵¹

(1) *The Nietzschean strategy* proceeds by rejecting *the normativity of order*. Consequently, the ethic of cognition is abandoned *tout court*. This abandonment manifests itself in the outright denial of human cognitive powers, rendering any proposed 'ethic' celebrating these powers superfluous. The idea of a cognitive grasp of an objective order achieved by science is replaced by the notion of science as one amongst the many perspectives conjured by humans in their ultimately lawless struggle for power. There is no 'truth' beyond these perspectives, no order to which they might correspond as 'true'. But the distinctiveness of the Nietzschean strategy lies not so much in its questioning the *possibility* of cognition - this it shares with all forms of scepticism - rather it lies in the challenge it makes to the *aspiration* towards truth. So when, for instance, Nietzsche famously describes truth - the norm of cognition - as "worn out metaphors which have become powerless to affect the senses"⁵², he is primarily contrasting and subordinating one norm *to another*. The challenge

to the norm of cognition has force in virtue of this contrast; the problem lies in specifying just what this other norm is supposed to be.

Nietzsche variously refers to it as "life" and "will to power". His suggestion is that our so-called truths are antagonistic to life, where "life" connotes a vital force of creative energy, a flux of sensuous particularity which resists the conceptual categorization which conditions claims to truth.⁵³ On Nietzsche's view, life expresses itself in metaphor - and more generally, in art. Nietzsche thus seems to be proposing a norm of 'expression', though not the expression of a subject about which there can be a truth, but rather of a deeper impersonal reality which manifests itself through the artistic creation of humans.⁵⁴ Although Nietzsche sometimes *conflates* truth with self-creation or "a will to overcome that has in itself no end",⁵⁵ his bizarre will to power metaphysics is adventitious to the challenge which opponents of his strategy must address; is it necessary to reject the norm of truth and the ethic of cognition for another norm that is external and *antagonistic* to it, and which must be thought of as having "in itself no end"? The force of this question turns on the relationship between rationality and the ethic of cognition. For if the two are equated, then a critique of the latter must invoke an "other of reason", not only because it is defined in terms of its antagonism to truth, but also because it has no other end - no other standard against which rational critique can begin.

The Nietzschean strategy, then, is to 'reconcile' cognition and identity by denying what is supposed to separate them; the very category of an order by appeal to which *either* cognition or identity can be secured. Not only is such an order unavailable, but the aspiration to pursue it is detrimental to proper intellectual conduct; self-creation and

self-transformation which has itself 'no other end'. In having 'no other end', the Nietzschean dispenses with the idea of a human identity *guided* by an order toward which the powers for self-creation may be properly or improperly directed.

(2) *Habermas's strategy* departs from Nietzsche's in attempting to preserve the differentiation of the claims of cognition and identity, while reconciling them on a formal rationalist ground. Habermas objects to the Nietzschean strategy on three counts. First, in hypostatizing the aesthetic moment of modernity, it is guilty of the same kind of one-sidedness which bedevilled Gellner's hypostatization of scientific cognition. Second, the appeal of this hypostatization presupposes the differentiation of cognitive from aesthetic claims which are simultaneously de-differentiated in the Nietzschean strategy. A similar kind of self-contradiction which Habermas thinks counts against it is that it cannot coherently be *claimed*, since in order to put it forward as a claim, its defender must presuppose norms built into the medium of raising and redeeming claims; what Habermas calls the communicative use of language. Indeed, it is just the rationality built into the procedure of communicative action which, for Habermas, makes available the moral order required for securing the identities of individuals and collectives. The Habermasian strategy is to argue for an order which is prior to and presupposed by the separation of cognition from identity: its point of departure is the shared medium of understanding between subjects capable of speech and action, or "linguistically generated intersubjectivity".

The normative structure of the medium which conditions cognitive claims is to be investigated by way of "rational reconstructions" of the implicitly known competences and rules presupposed in the achievement of mutual understanding. Accordingly, the

objectifying methodological attitude endorsed by the ethic of cognition loses its paradigmatic status, and is replaced by the "performative attitude of participants in interaction"⁵⁶. Self-understanding as rational reconstruction need not involve the objectification of self in the reflective gaze of a neutral observer - the kind of objectification which the Nietzschean strategy reacts against - for it proceeds by a "recapitulation" by the 'ego' of what comes into view from the second person perspective of the 'alter' in dialogue. And it is by reconstructing the non-coercive, unifying rational potential of agreements reached in the dialogue situation that Habermas aims to show that a conception of moral order without foundations in cognitive interpretations of nature but capable of sustaining the identities of individuals and collectives is available.

But the method of reconstruction cannot establish why, at the level of motivation, the moral order should be *followed*. To show this, Habermas turns to what he calls "therapeutic reason". He argues that there is a universal moral basis to modernity which when *distorted* generates instabilities and pathologies. He thus proposes a model of the distinctive pathologies of modernity under the presupposition that a standard is available for their critique. More specifically, Habermas argues that (a) the integrity of modern conditions of life is threatened by crises, manifesting themselves as pathologies, caused by the distorted reproduction of what he calls an "intact intersubjectivity"; and (b) that the norm of an intact intersubjectivity can be traced transcendently to the procedural conditions of the communicative use of propositionally differentiated language which is peculiar to the modern reproduction of life. Chapters three, four, and five elucidate and contend these claims.

(3) *Taylor's strategy* also departs from Nietzsche's in its commitment to establishing the *reason* for adopting it. For Taylor, the Nietzschean abandonment of order is irrational not so much because of the performative inconsistency involved in claiming it, since the Nietzschean can reply by disavowing any commitment to the canons of argumentation. Rather, the irrationality lies in its inarticulacy with respect to its alternative canon of normativity.⁵⁷ Additionally, it espouses a relativism which is incapable of making sense of the indubitable cognitive advances of modern natural science. Taylor's strategy seeks to avoid this Nietzschean anti-cognitivism, by way of making a conception of a moral order available which is cognizable but not by natural science. Whereas Habermas's strategy is to dissociate validity from its privilege in truth, Taylor's is to dissociate truth from its privilege in natural science.

Taylor's strategy for establishing this thesis is to argue for the ontological irreducibility of what is disclosed by human self-interpretations. The argument follows two steps; first, it aims to show that the concept of a human agent who prescind from horizons of morally significant self-interpretation is incoherent. This step is examined in chapter six. Second, it must establish that existential self-interpretations of this kind are equal candidates for cognitive value with interpretations of nature disclosed from the disengaged perspective of natural science, a claim which is considered in chapter seven. In the course of this argument, it becomes clear that as soon as the validation imperative is applied to humans, it too becomes one form of self-interpretation which must prove its superiority over others. The strategy is to take the status of the ethic of cognition seriously as an ethic - not in order to reject it a priori - but as the first step in showing its inferiority to other norms which confront it as rivals. And it is precisely because there are such rivals that the ethic of

cognition requires the kind of defence which Gellner offers for it.

But just because, as Gellner asserts, the cultural injunction called the ethic of cognition is so deeply ingrained in modern life, Taylor's strategy stands at a disadvantage. For it needs to be able to recover a sense of human identity which is occluded in a culture deeply shaped by disengaged cognition and technologically mediated practice. This means that the language through which the strategy is employed needs to have a special degree of resonance to be convincing. But there is an even greater handicap. For in such a culture, Taylor suggests, the *moral order* which is supposedly disclosed by language loses its foundation in the *public order* of references in the medium of which identities are initially forged, and in which knowledge claims are tested. But this conclusion threatens to undermine the coherence of Taylor's strategy for reconciling cognition and identity, since it dissolves the ground for thinking that linguistically disclosed moral and scientific understanding are ontologically symmetrical. This thought is pursued in chapter eight.

(2.5) Conclusion

To summarize: Gellner argues that the mechanistic mode of legitimating beliefs has a corroding effect upon the linguistic resources available for forging believable human identities. Science legitimates beliefs about the world according to formal explanatory criteria of public reproducibility in a minimal language of primary quality concepts. But this language is not enough to fill out a human identity, which means that modern human beings cannot legitimate beliefs about themselves in a cognitively satisfactory way. The imputation of a lifeworld, according to Gellner, expresses a failure of nerve in face of this predicament between cognition and identity. For it suggests a bracketing off of the world

as it is lived, a world which forms a background of belonging and source for the self-reassurance and legitimation of human identities, from the objective world which is the backdrop and reference point of genuinely cognitive claims. Now not only is this 'bracketing off' said to be a betrayal of cultural modernity's ethic of cognition which prioritizes truth above all other values, but it is also to misunderstand the relationship between science and modern everyday life. The bracketing does not work because it already presupposes an everyday world in which instrumental rationality, the formation of hypotheses to solve problems by discovering the maximally efficient means of bringing about a particular end, has a constitutive role. Insofar as the resource of the lifeworld is allowed cognitive status at all (Gellner's Fixed Cognitive Capital), it becomes destabilized by the continual growth of hypotheses concerning that status and of the instrumental rationality which can assess hypotheses (the so-called Variable Cognitive Capital). The outcome is a predicament because the ends by appeal to which instrumental rationality can assess value - the claims of the lifeworld - have in this rationalization process lost their immunized status, without thereby acquiring a rational one. They appear to be relativized, arbitrary, optional, whilst identities and ends of action as such can be none of these things. I have argued, however, that by dismissing the lifeworld concept, and by prejudicing the case against forms of rationality other than instrumental efficiency and mechanistic scientific validation, Gellner fails to give good grounds for accepting the diagnostic thesis which concludes his argument. He unwittingly disposes of the very conceptual resources he needs to give it a convincing formulation. The onus is now back on Habermas and Taylor to show how identities and ends of action - or the claim to normativity - can be secured from the fate of having a merely optional, hypothetical status.

CHAPTER THREE: HABERMAS'S CONCEPTION OF CRISIS

(3.0) Introduction

In the previous chapter, I introduced two different strategies for responding to a tension between identity and cognition wrought by modern scientific practices of legitimating belief. Habermas shares Gellner's commitment to the view that in the course of its evolution, the human species comes to learn that the nomological order of nature is empty of moral signification. It can then seem impelling for Habermas to take on the guiding problematic of Kantian philosophy: how is the source of moral significance to be grounded if not in the natural order of *things*? Aware of the threat which the objectivity of modern natural science posed to the identity of the reflective human subject - a reflexivity which is *imposed* upon moderns with the collapse of the social fabric of religious tradition - Kant sought to bring the ground of moral order to self-clarification in the rationality of free and responsible action. There may not be any ethical substance in the world which can be the object of cognition, but there is a moral law which can be tracked by a free subject acting on universalizable self-willed maxims. And in the moral law, the human agent can be reassured that there are reasons for acting which have more than a merely optional, hypothetical status, since there is a class of actions which, *qua* rational beings, humans have an unconditional obligation to perform. The moral law, for Kant, puts constraints on matter-of-fact, empirically motivated action by opposing to the particularity of sensuous inclination the universality of the rational, dutiful will. For Kant, the struggle for normativity can only mean the attempt to institutionalize the moral law as it is revealed to the conscience of each rational reflective subject.

But by abstracting the rational agent from the historically concrete intersubjective conditions of agency-formation, and similarly by idealizing the source of the legitimacy of the moral law as *prior* to any institutionalized embodiment - Hegel famously countered - the Kantian view fails to give a plausible account of just what is threatened by the objectivating sciences and the fragmented social fabric of religion; the lived sustainability of a moral identity. Such an account can only be given, in the view proposed by Hegel in his early writings, if instead of understanding the moral order as what is categorically determined by the moral law, it is conceived as an 'ethical totality' which exercises its compulsion upon acts which transgress it as a 'fate'.

Now it is just as a 'causality of fate' that Habermas conceives the systematic dynamics of moral crisis and social pathology formation in modernity. "The pathological characteristics of modern societies now fit into patterns", Habermas writes, "only to the extent that a predominance of economic and bureaucratic rationality - of cognitive/instrumental forms of rationality generally - makes itself felt".¹ But this predominance does not make itself felt *directly*; it requires theoretical self-clarification. The avowed achievement of Habermas's Critical Theory is to do this by way of reconstructing "Hegel's concept of the ethical context of life" in such a way that it "disenchants the unfathomable causality of fate".² The characteristic crises of modernity are explicable, Habermas puts it otherwise, as a struggle for normativity which can be brought to philosophical clarification in terms of a "dirempted totality, which makes itself felt primarily in the avenging power of destroyed reciprocities and in the fateful causality of distorted communicative relationships".³ It is through this "avenging power" that the force of normativity, now taken as the force of the rationality potential of undistorted communicative intersubjectivity, makes itself felt in conditions of

a predominant "norm-free" cognitive/instrumental rationality. According to the model Habermas claims to borrow from Hegel, a moral order discloses itself behind the backs of (individual or collective) human agents *in* the identity crises they suffer as a result of disturbed self-formative processes; processes which refer back to an immanent ethical totality. In this chapter, I want to consider the kind of theory Habermas takes to be best equipped to articulate this model of crisis and its reflective overcoming. For it is from such a model that Habermas extracts the conception of a moral order which can fill the identity-gap left by religion and hence provide modernity with self-reassurance.

In (3.1), I outline how the notions of an ethical totality and causality of fate feature in the young Hegel's critique of the Kantian construal of a moral order as the moral law. Against Kant, Hegel proposes a tragic model of the self-formative process of the moral subject by appeal to the idea of a fatefully avenging ethical totality. At this point, for the sake of clarifying Hegel's insights, I compare them with views recently espoused by Bernard Williams. Following this, I introduce Habermas's incorporation of the young Hegel's critique of Kant into his general diagnostic thesis of crisis-formation in modernity as a struggle for normativity. This general pattern of disturbed and reconstituted moral identity-formation, Habermas once claimed,⁴ forms the implicit basis of what is metatheoretically interesting and correct in the diagnostic and therapeutic claims of Freudian psychoanalysis. The issue of what *kind* of theory is capable of articulating Hegel's insights is taken up in (3.2). Here I discuss the metatheoretical implications Habermas draws from understanding the kind of psychopathologies investigated by Freud according to the Hegelian model.

I then consider two kinds of objection which have been launched against Habermas's synthesis of Hegel and Freud; first (3.3), that in its neglect of the scientifically discoverable laws of human development, it is insufficiently naturalistic; and second (3.4), that in pandering too much to the requirements of an empirical science and to the injunction of cognitive/instrumental rationality, it is too naturalistic to serve its putative 'critical' purpose. I distinguish two different meanings of the 'disenchantment' of the causality of fate which Habermas claims to be the achievement of his preferred kind of critical theory, anticipating objections I will put in chapter five concerning Habermas's equivocal use of the concept of 'ethical totality'. Since the normative foundations of Habermas's critical theory are tailored to his conception of crisis - and hence the problem of self-reassurance - Habermas will also need to clarify how the normative presupposition of an ethical totality can be articulated in a theory which meets the two kinds of objections considered in (3.3) and (3.4). That move will be considered in the next chapter.

(3.1) Ethical Totality and the Causality of Fate

As is clear from the remarks cited in (3.0), Habermas's appropriation of Hegel's model of an ethical totality which avenges itself as a fate is decisive for his proposed resolution of the problem of self-reassurance. It is important, then, that some consideration be given to the purpose to which Hegel originally employed these concepts. In *The Spirit of Christianity and its Fate*, the young Hegel can be read as offering a phenomenology of morals by way of describing two contrasting appearances of the binding force of the moral relation.⁵ Moral authority, it is assumed, has the character of universality, inasmuch as particular agents stand to it in a relation of some kind of compulsion or guidedness. Hegel attempts to make sense of this phenomenon of the binding, universal, authoritative (as we

would now say, 'objective') nature of morals, from the point of view of a particular subject's lived experience of its compelling force. From this point of view, the issue emerges of the possibility of *reconciliation* between the compelled particular, and the compelling moral source. This is an issue which Hegel can focus on by considering the structure of the lived experience of moral transgression; of crime and punishment. The criminal experiences the compelling character of moral authority, which he has usurped in his act, as punishment. But the meaning of the punishment will differ depending upon whether it appears as the revenge of law, or as the avenging force of what Hegel calls 'fate'. Although in Hegel's text the same term 'punishment' is used to refer to both 'law-like' and 'fate-like' kinds of avenging force, I shall hereafter demarcate them by capitalizing the 'P' of the former kind. So where the binding force of morality appears as the force of the law, a criminal act brings into play the avenging force of Punishment.

The Punishment suffered by the criminal would appear under the following characteristics. First, Hegel describes it as the necessary and inescapable consequence of the criminal act: Punishment is *entailed* by the act. If Punishment represents the avenging force of the moral source (the source of normativity), and if the moral source is represented by the moral law, then the 'must' of the Punishment will be as the imperative of the moral law - namely, categorical. The necessity of the Punishment reflects the imperative of the law. As the imperative of the moral law is necessarily and unconditionally (i.e. categorically) applicable to the human being *qua* agent, so Punishment is applicable unconditionally and necessarily to the agent *qua* criminal. This, as Hegel sees it, is the unbending demand of justice; "so long as laws are supreme, so long as there is no escape from them, so long must the individual be sacrificed to the universal" - i.e. the law, in Punishment.⁶ But it is only

possible for the individual to be sacrificed to the universal in this manner, thinks Hegel, if the individual is *identified* with his standing with respect to the law. That is, with respect to the law, the criminal's identity is *nothing* but that of the perpetrator of crime. From the moral point of view - the point of view of the judge - he *is* a criminal; "a sin existent, a trespass possessed of personality".⁷ Now the criminal would be *only* that, from the moral point of view, if the law felicitously represented the whole of the moral source. Yet it is an identity which an individual can acquire only by being abstracted from the concrete conditions of his life context. Thus Punishment also appears to the criminal as exercising its avenging force upon an abstraction. In Punishment, the agent who trespasses against the law is considered only in the abstract. Third, and crucially, the moral source which exercises its avenging force as Punishment appears as *external* to the criminal. The law is external to the trespass, it stands outside and is unchanged by the transgressive act. The fear which the criminal has of the universal is a fear of something which stands above and beyond him - as something alien. This is the case, Hegel suggests, even if the criminal would have willed that law himself as a universal maxim of action. For although qua maxim the law may be self-willed, its reactive force (suffered as Punishment or guilt) stands independent and external to that will. The transgression of his duty appears not only as an unworthiness, but as the misfortune of a concrete being with particular inclinations living in circumstances not wholly under his control.⁸

I indicated that in making a distinction between the avenging force of law and the avenging force of fate, Hegel is especially interested in the possibility of reconciliation between the particular person who acts, and the reactive universal moral force which appears to the agent as following ineluctably from the act. The three characteristics of Punishment just

outlined do not readily accommodate this possibility. First, the criminal faces the necessity of Punishment, so far as the law will have its way. As a matter of fact, of course, the law may not have its way; any actual judge is a concrete individual with particular feelings and inclinations which - though contingent to his station - may affect his judgement (the criminal always has the chance of getting away with it). But as Hegel puts it, this "contradiction between consciousness of oneself and the hoped-for difference in another's idea of one's self", or the "contradiction between desert in the eyes of the law and the actualisation of the same",⁹ is no basis for a *reconciliation* with the law. Second, given the abstraction of the criminal qua criminal from the point of view of the law, any putative reconciliation with the law could only be what Hegel calls a mere "conceptual reunion", involving "man as a concept" rather than "man as reality". And finally, as long as crime appears as the "destruction or subjugation of something alien", external to and unchanged by the act, again the thought of reconciliation is absurd.

Hegel's conclusion is that within a framework structured by unmediated oppositions between 'universal' and 'particular', 'concept' and 'reality'; "Punishment and law cannot be reconciled".¹⁰ However, he wants to show that reconciliation is possible if these oppositions can be transcended. In what is all but an echo of the young Hegel's fragment, Bernard Williams has expressed the need for such a transcendence with impressive economy.¹¹ The Punishment through which the law avenges itself corresponds to what Williams calls "the institution of blame" which bears upon the transgressor of the peculiar "system of morality".¹² Within the system of morality, the fundament of ethical life is construed as an unconditional obligation to act in accord with the moral law. The phenomenal appearance of an objective constraint on a subject's particular inclinations is

then taken to represent an implicit recognition of the universality of the moral demand - as an intuited sense of one's overriding obligation to a moral law which is universal in virtue of being uncompromised by such inclinations.¹³ This is taken as the achievement of the rational will. But as Williams describes it, the experience of the moral demand is like being confronted with something which is "part of the world in which one lives". The force of the moral law is explicable as external to the individual rational will only if it is regarded as the law of a "*notional republic*", in which the law is rationally self-imposed by each citizen.¹⁴ The morally construed law-like fundament of ethical life, in Hegel's terms, impinges on 'man as concept', and it does so by abstracting the demands of law from the concrete world in which 'man as reality' lives.

Moreover, the *stability* of the morality system, according to Williams, requires a means of *binding* individuals which it is itself incapable of providing. The system of morality is inherently unstable, Williams suggests, because it does violence to the reality of the individual's life context. In order to be grounded as law, the system of morality needs its own peculiar sanction, and this it finds in a particular kind of punishment - in blame. Allocations of blame tend to focus narrowly on an action or omission perpetrated by an isolated, ideally autonomous subject. It can then seem as if the agent who transgresses the system of morality always has a reason for acting otherwise - namely, to fulfil his moral obligation. But this reason is then viewed as overridden by another reason upon which the transgressor chose to act, and for which he merits blame. The practice of blame thus presupposes that there is a common (or 'basic') reason for acting which all agents implicitly or explicitly recognize, and to which the agent can voluntarily conform or dissent. Not only is such a basic reason fictional, Williams argues, but so is the idea of the ideally

autonomous subject which chooses to act upon it. The practice of blame abstracts moral consciousness from the background context of self-formation, and so from the surroundings in which the shape of the particular character blamed is forged. At this point, Williams suggests that the fictional foundation of the system of morality and its misrepresentation of the pull of the ethical as blame can be overlooked if one is convinced by the ideal that "human existence can be ultimately just"¹⁵; that luck in the process of character-formation and the leading of life are adventitious to what is of fundamental value in life. Williams notes that this aspiration towards 'purity' may have been of some beneficial practical consequence. More sceptically, however - and here the echo of Hegel is at its clearest - Williams writes that the fiction underlying the practice of blame can encourage people "to misunderstand their own fear and resentment... as the voice of the Law".¹⁶ Williams concludes that although the fiction of the moral law may have augmented the amount of actual justice in the world, it does so by misrepresenting the nature of the limited value it fosters. This misrepresentation can only be overcome - and the end of justice more felicitously served - if the illusion of value without luck is dropped, and with it the irreconcilable oppositions between duty and inclination, voluntariness and force, and with them the idea that unless social practices embody justice purely, they must be failing to institutionalize justice at all.¹⁷

But where does the acknowledgement of luck leave us, and how might these oppositions be reconciled?¹⁸ Let us now consider the reconciling or redeeming potential of what Hegel calls the avenging force of punishment as fate. We saw that for Hegel, the avenging force of the law as Punishment takes the form of a sacrifice of the individual agent as an abstract category to the universal demands of morality, the necessarily hostile force of which is

fearfully experienced as alien to the particular individual. Taking these features as considered above in reverse order, the appearance to the criminal of his punishment as a fate contrasts with them, according to Hegel, in the following ways.

First, what avenges itself in the power of fate is not something external to the agent, but the 'defective life' of the agent *himself* insofar as this life is shared with the other members of his ethical community:

When the trespasser feels the disruption of his own life (suffers punishment) or knows himself (in bad conscience), then the working of his fate commences, and this feeling of a life disrupted must become a longing for what has been lost. The deficiency is recognized as part of himself, as what was to have been in him and is not.¹⁹

What has been lost is not the satisfaction of his own particular self-interest as sacrificed or opposed to a universal moral law. Hegel speaks of what is destroyed by the trespass as the "friendliness of life". The longing for what has been lost is the longing for the friendly context of life which through his *own* act, the criminal has disrupted, or distorted, into an enemy. Hence, the fear of the criminal is directed not towards something which pre-existed his act and which is external to it. The criminal's fear of punishment as fate is a "fear of a separation, an awe of *one's self*"²⁰; he makes himself into his own enemy by separating himself from, and thereby making an enemy of, what was to have been but is not a friendly context or totality of life. For this reason, the moral significance of his act is not primarily the annulment of an external moral law, but the *diremption* of a shared ethical totality.

Now if in his crime the criminal becomes 'split-off' from the presupposed (in the longing)

friendly context of life considered as a *totality*, it follows that *all other* parties to the totality suffer the consequence. Not only the criminal, but all other members of the ethical community are implicated. The avenging force of punishment is thus fateful in the following sense; it draws in all parties to the moral relation, even those who are innocent of any crime. Hegel's point is that the criminal's act is not in itself sufficient to bring the causality of fate into play (as in the force of law). Hegel distinguishes between the act as the occasion of this force, and the reaction to it on the part of the other members of the moral relation which produces it. The injured party can either struggle for the recovery of his right, or renounce it in "submissive grief".²¹ In either case, though neither doing wrong nor deserving punishment, the other finds himself in a position of *responsibility* "as an inescapable fate".²²

The necessity of this fate, therefore, is not to be considered in terms of what is entailed by the abstract identity of an agent qua criminal from the point of view of the law. For not only is the fate of the criminal not just his, but the reason for this being so is unthinkable outside the concrete conditions of his life context. And it is the avenging force of this distorted concrete life context which the criminal experiences in his punishment as fate. This is not to say, to return to the first of Hegel's contrasts, that justice is compromised by fate, since "even in the hostility of fate a man has a sense of just punishment".²³ The point is that this hostility is not *grounded* in the moral law, but in the ethical totality - the friendly context of life - presupposed in the appearance of a dirempted life which avenges itself as a fate. The principle of justice which is the unbending demand of the moral law is but a fragment, albeit an important one of the ethical life. Since fate "knows no precinct of virtue"²⁴, we can say that for Hegel, Punishment presupposes punishment; "The law is

later than life and outflanked by it", "the law is only the lack of life, defective life appearing as a power".²⁵

This power is what appears to the criminal as the compelling force of moral authority. But the universal it represents, Hegel insists, "is not severed from the particular in the way in which the law, as universal, is opposed to man or his inclinations as a particular".²⁶ For in the source of the compulsion, the criminal recognizes his own life as a particular individual as it might have been before his separation from it in his punishment. Referring to this opposition between universal and particular, Hegel makes the point that;

Before he acts there is no cleavage, no opposition, much less a mastery. Only through a departure from that united life which is neither regulated by law nor at variance with law, only through the killing of life, is something alien produced. Destruction of life is not the nullification of life but its diremption, and the destruction consists in its transformation into an enemy.²⁷

If the universal of moral authority is considered in terms of the ethical totality - or friendly context of life - reconciliation becomes possible between it and the particular agent who through his crime, becomes split-off from it. For Hegel, reconciliation between split-off or divided fragments of life, and the return to the friendly life which has been lost, is the achievement of love. In love, as he puts it, "fate is reconciled".²⁸

Now Habermas has always acknowledged the force of the young Hegel's exposition of the weaknesses in the Kantian conception of the basis of morality.²⁹ Particularly, Habermas identifies those inadequacies which result from Kant's abstraction of moral action from the particular inclinations which motivate it; of the form of duty from the specific,

context-dependent content of its application; but most importantly, of the autonomous will from the complex intersubjective nexus of its *self-formation*. And it is precisely the emergence of moral identity out of a destroyed intersubjective nexus of mutual recognition which Habermas takes to be the central insight of Hegel's position. "In the causality of destiny", Habermas writes,

the power of suppressed life is at work, which can only be reconciled, when, out of the experience of the negativity of a sundered life, the longing for that which has been lost arises and necessitates identifying one's own denied identity in the alien existence one fights against. Then both parties recognize the hardened positions taken against each other to be the result of a separation, the abstraction from the common interconnection of their lives - and within this, in the dialogic relationship of recognizing oneself in the other, they experience the common basis of their existence.³⁰

It is this common basis of existence which Habermas takes to be the source of morality. It is the goal of that struggle for normativity the self-clarification of which is the task of his critical theory. He describes it as "the complementary interchange of non-compulsory communication and the mutual satisfaction of interests".³¹ Habermas's claim is that this structure of intersubjectivity is a necessary condition of undamaged self-formative processes. It thus provides a standard which can be appealed to in the critique of the systematic injury to human identity which makes itself felt in the predominance of cognitive/instrumental reason. But Habermas *also* wants to claim that it is within this undamaged process of self-formation - or 'intact intersubjectivity' - that communicative *rationality* is deposited. As Habermas puts it in a more recent formulation;

Any violation of the structures of rational life together, to which all lay claim, affects everyone equally. This is what the young Hegel meant by the ethical totality which is disrupted by the deed of the criminal and that can only be restored by insight into the indivisibility of suffering due to alienation.³²

The problem Habermas now faces is that of making Hegel's insight intelligible within a general account of the emergence of structures of rational life. But given the theoretical commitments already introduced, this can only mean within an account of the evolution of communicative competence as organized within a *theory of rationalization*. So Habermas must find a way of combining the conception of crisis and the revelation of a moral order which he derives from Hegel with the Enlightenment derived conception of modernity as guided by the norms of procedural rationality. It is not initially clear how such a synthesis can be forged, since procedural reason departs from its substantive, pre-modern precursor just in its differentiation of truth-evaluable action-motivating moral insight from the self-clarification of the claim to normativity; where the achievement of this separation is a measure of communicative competence. By seeking to integrate Hegel's insight with a theory of rationalization, Habermas thus faces the danger re-introducing the Kantian abstractions against which Hegel's critique was originally directed. In the general guiding terms of our discussion; are the acultural conceptual tools needed to theorize modernity as a process of rationalization suitable for articulating the injuries inflicted upon an ethical totality as a fate? I now want to turn to how Habermas takes the causality of fate to be evidenced in actual self-formative processes, and to the kind of theory which Habermas claims can best articulate its operation. This is most clear in Habermas's interpretation of the achievements of Freudian psychoanalysis.³³

(3.2) The Fateful Causality of Psychopathologies

Habermas takes as his basic model for the conflicts of modernity which generate its need for self-reassurance "the pseudo-natural dynamics of impaired communicative life-contexts".³⁴ The notion of an unimpaired communicative life-context is the standard against which these conflicts are to be understood as symptoms of pathologies. At this point I want to consider the pattern of *dynamics* Habermas attributes to pathologies arising from the impairment of this standard. For these dynamics are *pseudo-natural* precisely to the extent that a causality of fate appears as a causality of nature. The distinction between these two dynamics is most fully elaborated in Habermas's discussion of Freud in *Knowledge and Human Interests*,³⁵ where he presents an argument to the following effect. If Freud was theoretically correct in inferring a dynamic of inner conflict from phenomena of resistance within a subject's psyche to part of its own content, and if Freud was therapeutically successful in undoing this resistance through an act of communicated recollection which was the subject's own doing, then the dynamic of the pathological inner conflict must be articulable by the critical concepts of self-recognition and responsibility. But since self-recognition is possible only within a public context of linguistically shared rules of interpretation, the private dynamic of inner conflict, resistance, pathological self-misrecognition and emancipation through self-responsible reflection, will refer to a public dynamic of a distorted and reconstituted *grammar*. For Freud, the patient's pathology is essentially determined by the law-like causality of the instincts, such that his suffering before his symptoms can be compared to what Hegel described as the suffering of the criminal before an alien, unbending penal law. But for Habermas, the patient stands to his symptoms as the criminal described by Hegel stands not to a law but to his fate. The common fate of the patient and the criminal is the suffering of an excommunication, which

for Habermas means an alienation not from an external moral or libidinal law, but from just that *communicative* life-context in which processes of self-formation occur. Distorted self-formation thus reflects on what Habermas is calling an impaired communicative life-context, or damaged intersubjectivity.

I shall now investigate this argument in more detail. The distortion of communicative life-contexts is Habermas's model for the social pathologies characteristic of modernity. According to Habermas, this is just the sense in which pathologies were identified and treated by Freud, though Freud's bewitchment by the model of the natural sciences led him to misunderstand this. The issue I want to focus on is not the validity of Habermas's epistemological critique of Freud's science, neither is it the fidelity of Habermas's interpretation of Freud's system as whole. Rather, my main objective is to elucidate the sense in which psychoanalytical theory is concerned with articulating the operation of a causality of fate in processes of pathological self-formation. For this is the same dynamic which, according to Habermas, generates the crises in identity he elsewhere attributes to the distorted reproduction of modernity's symbolic resources.

"The starting point of psychoanalytical theory", for Habermas's purposes, "is the blocking force that stands in the way of free and public communication" of a class of psychic contents.³⁶ In everyday life, we often refrain from freely and publicly communicating episodes in our mental life; one may feel conscious reservation about communicating something which is compromising, tactless, or unedifying either to one's partner in communication, or more commonly, to oneself. Words can hurt, and aware of the social and moral pressures to avoid hurt, we can *choose* not to communicate such words.³⁷ But

in the situation of dialogue between analyst and patient, these normal everyday restraints to communication are bracketed. We have an 'experimental' situation in which the 'initial conditions' are such as to maximize the extent of freedom on the part of the patient to publicly communicate his or her mental history to the analyst. Psychoanalytical theory *begins* with experiences of resistance here. In the situation of dialogue between analyst and patient, empirical data in the form of amnesias in the recounted mental history of the patient are interpreted as evidence of resistance within the patient to key episodes in that history. The two-fold task of psychoanalytical technique is to overcome this resistance and to interpret what is latent to the manifestations of it. Overcoming resistance involves neutralizing the blocking force to the communication of contents of the patient's psyche which have become 'lost' to him. The nature of the amnesias encountered by the analyst suggests that it is not simply the passive loss of memory which is at work here - a temporary excursion into the resting place of the pre-conscious - but a loss which is due to an active force of resistance which works 'behind the back' of the patient. The resistance experienced by the analyst in his or her effort to get the patient to recall an episode in the patient's mental history - under such conditions - reflects on the unconscious resistance experienced by the patient him/herself to the very same phenomenon.

Freudian psychoanalysis attempts to capture the logic of this blocking force in the concept of repression, where the class of psychic contents which resists open and public communication is that content of the mind which is systematically resistant to consciousness. This psychic content is both affective and symbolic. It includes needs, desires, and motivations, which at their most thoroughgoingly and elementally affective level, are referred to by Freud as instincts. They are, so to speak, the bare substance of

the psyche considered dispositionally. However, insofar as we can talk about an affective content of the *mind*, this content must be given form by being 'attached' to the specific means of operation of the mind. As the mind operates by way of symbols articulated in a language, so the affective content of the mind must be symbolically or linguistically mediated. To have an intelligible mental content is to have acquired a language. But to have acquired a language, Habermas learns from Wittgenstein, is to have been introduced into, and to participate in, a linguistic practice or 'language game'.³⁸ This latter notion contains the idea of language as a rule-governed activity of communication, a necessary condition of which is that the rules which make the communication of meanings possible are public. This is not because meanings are initially private to the mind of a subject, and are then contingently made public in order to communicate them to an other mind. Rather, meanings are constitutively public, and are only understandable to oneself for the very same reason that they are understandable to others, publicly, in the act of communication.

But if psychic contents are linguistic in constitution, and language is constitutively public, what bearing does this have on a 'repressed' content of the mind? As Habermas interprets it, the repressed content of the mind refers to just those linguistic interpretations of "need dispositions" which are excluded from public communication. Since social pressures may prohibit the communication of certain motivations or need dispositions, the symbols through which they are interpreted can become 'split-off' from the public rules which are the grammar of "the ongoing text of our everyday language games".³⁹ But since one understands one's *own* motives, actions, and patterns of expression according to these public rules, then those motives, actions and patterns of expression the interpretations of which are removed from public circulation appear incomprehensible to the subject or author

of them. And this, according to Habermas, is precisely the incomprehensibility of the unconscious. The unconscious thus corresponds to the repressed content of the mind, where what is repressed is a public language for the interpretation and communication of need dispositions which are sanctioned against socially. Hence the resistance which is the starting point of psychoanalytical theory - namely, the resistance to free and public communication of repressed psychic contents - can now be seen to correspond to the resistance of the unconscious content of the mind to consciousness.

The problem of comprehending the unconscious is the second of the two-fold task of psychoanalytical technique; to interpret the manifestations of resistance. For although an event in the subject's life-history may be displaced inaccessibly from her consciousness, the suppressed need-disposition it represents nevertheless continues to leave its trace. "Because the symbols that interpret suppressed needs are excluded from public communication", Habermas writes, "*the speaking and acting subject's communication with himself is interrupted.*"⁴⁰ This interruption in the communicatively mediated identity-formation of the subject has considerable repercussions upon the apparent expression of that identity. And it is through the interpretation of these repercussions that the 'lost' content may be retrieved. In extreme cases - such as neuroses - they manifest themselves publicly as pathologies in the form of 'symptoms'. Following Freud, Habermas theorizes neurotic symptoms as a defence mechanism of substitute gratification for early experiences of traumatically suppressed needs. But more distinctively Habermasian is the view that symptoms, such as compulsive repetitions, phobias, and obsessional fixations, are distorted expressions of dispositions the interpretation of which has been excommunicated from the public realm. The subject does not understand why he or she repetitively and compulsively

acts in ways which defy his or her conscious preferences and beliefs, and awareness of this discrepancy is of no effect in changing the actions to conform with these preferences. Symptoms are actions which belie the subject's motivations, the latter appearing to have a compulsive force which is external to the agent. So the interpretive problem can now be specified as that of making sense of; (1) the belied motivation, in terms of; (2) an excommunication from the public realm of meanings, which; (3) exercises an externally compelling force. It is this third aspect of the problem which most interests us, so I shall just very briefly deal with the first two aspects of the so-called 'depth-interpretation'.

(1) The symptoms and conscious intentions of the subject belie the dispositions which motivate them in the sense that the latter are expressed in a distorted, manifestly disguised form. They appear as *irrational*.⁴¹ But since both the (verbal and non-verbal) conscious actions and the unconscious dispositions are parts of the same 'self', then Habermas considers the belying of one by the other to be a form of self-deception. Because of this, psychoanalytical interpretation takes on a peculiar character. Its interpretive dimension invites comparison with hermeneutics, the goal of which, as we saw in (1.3), is to interpret the problematic meaning of historically and culturally distant texts. Likewise, symptoms (and in normal cases, parapraxes, dreams, etc.) can be considered as texts the elusive meaning of which requires artful interpretation. But in contrast to hermeneutics, psychoanalytical interpretation is directed towards non-accidental textual distortion; "The omissions and distortions that it rectifies have a systematic role and function. For the symbolic structures that psychoanalysis seeks to comprehend are corrupted by the impact of *internal* conditions."⁴² That is, they are the product of 'systematically distorted communication' the comprehension of which requires tools drawn from a general theory.

of communicative competence which transcends the domain of everyday, theoretically unmediated hermeneutic consciousness.⁴³

Psychoanalysis seeks to uncover intentions which have in the sense described become 'unconscious'. The meaning of the symbolic structures (texts) are not transparent to the subject (author), and the goal of interpretation is to make these meanings accessible to the subject. The interpretation can meet this goal, Habermas believes, by simultaneously *explaining* how the inaccessibility, or distorted meaning, has come about. The flaws in the text of the actions and recounted mental history of the patient, the symptoms and the amnesias, are not arbitrary distortions of the intentions or dispositions of the author. They have meaning as such *as* resistances. The self-understanding generated in the analysis "makes accessible the meaning of specifically incomprehensible forms of expression only to the extent to which it is possible to clarify the conditions for the emergence of non-sense in conjunction with the original scene".⁴⁴ But if the distortions are not accidental, what is the nature of the causality or necessity operating here, and how can it be reversed? Before attempting to answer this, we must briefly attend to the second aspect of the interpretive problem.

(2) The analyst notices a lack of 'fit' between what is linguistically stated and what is non-verbally (bodily) expressed by the patient. They apparently make no sense together. But if symptoms, as Habermas has claimed, are distortions arising from a splitting-off of the need disposition from a publicly meaningful language, then the peculiar task of psychoanalytical interpretation is that of a translation from the privatized or 'degrammatized' text of the symptom (or in normal cases, the dream) to the public language

of linguistically storable intentions. The interpretive task involves "translating symbols from a mode of expression deformed as a private language into the mode of expression of public communication", thus bringing "to consciousness the person's own self-formative process".⁴⁵ Translation is possible precisely because the distortions of the subject's text are not arbitrary. The distortions are "meaningful as such", in that they both resist conscious (public) articulation, while at the same time revealing this resistance in disguise; "subjects deceive themselves about themselves through language and simultaneously give themselves away in it".⁴⁶ To unmask this disguise, the making manifest of the latent content of patient's psyche, is to articulate this content in a language which is accessible and meaningful to the patient, which is the public language in which each person's self-formation is comprehensible. That is, in the grammar of the ongoing text of everyday language-games.

The position reached so far is this. Starting from Freud's insight into the role of resistance in freely and publicly communicating aspects of one's psychic history, Habermas has reformulated the concept of repression in terms of an excommunication, "carried out in and through language"⁴⁷, of aspects of one's self-formation. He can now say that the causality at play in repression and symptom formation is the kind of causality which is appropriate to the process of self-formation. But it is a distorted or pathological process because, as a result of repression, one comes by patterns of behaviour and bodily expression which are alien to that self, thus frustrating the satisfaction of desires which the subject would otherwise *choose*. Habermas uses Freud's phrase of "internal foreign territory" to describe the phenomenon of symptoms. They are foreign because they are not recognizable as the subject's own, they need translation into a language the subject understands. Yet they are

internal because they are also an integral, though fragmented, expression of the subject. Habermas must now indicate what the kind of causality appropriate to this process is, and thereby to show how *reconciliation* is possible between the alienated aspects of the self generated through it.

(3) The phenomenon which needs to be accounted for is the *externally* compelling character of the unconscious. The force of the compulsion of the unconscious is experienced as external to the subject because it manifests itself in actions and dispositions which are incomprehensible to the subject, yet these actions and dispositions are nevertheless part of that subject's identity. This externality of the compulsion tempted Freud to attribute the causality of the unconscious to natural laws. The physical forces which govern the behaviour of the unconscious according to laws of nature Freud called instincts. Habermas in effect makes three objections to this move. First, he claims that the concept of instinct is inapplicable to mental contents. The very concept of instinct is only applicable at all, so he claims, because animal behaviour can be understood in terms of a reduced, but irreducibly linguistically interpreted human experience, of hunger, love, and aggression. It is inapplicable to the mental content of humans because it is divorced from the specifically linguistic character of this content. In this connection, Habermas opposes a pseudo-natural causality of instincts to a causality of fate which "prevails through the symbolic means of the mind".⁴⁸

Put this way, Habermas's idea of a causality of fate seems to be close to the view that symbolic representations which are reasons can be causes of actions without nomologically determining them. The causality which prevails through the symbolic means of the mind

is, according to this interpretation, the causality of reasons, which don't appear to be connected with particular actions in a law-like way. In particular, the connection between a symptomatic action and the initial desire and defensive reaction cannot be satisfactorily explained in terms of the invariance of natural laws. But there is invariance of *some* kind; what Habermas calls "the spontaneously generated invariance of life-history" ⁴⁹, which is captured in Hegel's notion of a causality of fate. Habermas's second objection, then, is that it is not necessary to invoke the concept of instinct to account for the causal conditions of the unconscious, since an alternative kind of causality to the law-like causality of "natural" conditions is available to us. This is the causality of reasons for action which have become 'split-off' from the communicative life-context of self-formation. According to Habermas's 'excommunication' model of the unconscious I have been outlining, the repressed content of the mind (the unconscious) appears only with the banishment or exile of unconstrained and publicly communicable need interpretations and expressions. The reactive force of the excommunicated reasons for acting then appears as something law-like and external - as a 'second nature'. The subject's actions then appear to be pulled by this "objectified unconscious" as, say, the tides are by the gravitational pull of the moon.

This brings us to the third, and crucial objection which Habermas makes against the natural law-like compulsion of instincts - that it is not compatible with the power that self-reflection has to overcome the pathologies caused by repression and the causality of the unconscious. Habermas's transcendental argument is that the capacity for dissolution of the pathology by the self-reflection of the subject presupposes that a causality of fate rather than nature is at play in the original formation of the pathology.⁵⁰ If the latter were the case, Habermas argues, then it would make sense to consider the therapy as an essentially mechanical

procedure, however difficult, which could in principle be *repeated* in all like cases to similar effect. But the therapeutic achievement of an analysis is not analogous to the technical achievement of a biochemist in removing a pathological formation from a diseased or maldeveloped organism. For the achievement in the former case requires an act of *self-reflection* by the patient, and for this, Habermas insists, there can be no technical substitute. On the contrary, it is just in the emancipation of the subject from the 'objectification' which appears as a law-binding external second nature that the goal of the therapy lies. The act of reflection brings a practically momentous 'enlightenment' by freeing the subject from the occupying forces of the internal foreign territory. Reflection brings about a *reconciliation* between the conscious self and those symbols and motives which have "gone underground" and which appear as alien to the subject, while belonging to him. Self-reflection "transforms the pathological state of compulsion and self-deception into the state of superseded conflict and reconciliation with excommunicated language".⁵¹

So for Habermas, the 'cure' of analytical reflective insight is intelligible only if the 'cause' is attributable not to instincts exercising their force externally to the self, but to a division or diremption within the self which is reconciled by it. It works by 'undoing' the original process of splitting-off, and thereby reconstituting a "*grammatical* connection between symbols"⁵². But what allows this reversal to come about? What is special about the recollective insight? As Habermas reads Freud, the answer lies in the peculiarity given by its contextualization in the 'transference' situation. During the course of an analysis, the patient can come to transfer his or her reaction to the symptom-forming occasion from the original scene of the symptom, to the artificially controlled analysis situation. Analytical self-reflection renders the latent meaning behind the apparently irrational and

misery-inducing symptom comprehensible "by reference to the unmutated meaning of the original scene in infancy".⁵³ This "scenic understanding" is rendered possible under the presupposition that an equivalence is established between the everyday symptomatic scene, the transference scene, and the original scene. With the pressures and social obligations of life suspended by the analyst, the patient's symptoms are allowed unconstrained expression, thus weakening the resistance of the patient to recollecting the original scene. With the help of the reconstructions of the original scene suggested by the analyst, and "confronted with the results of his action in transference", the patient can come to recall the lost portion of his or her mental history seen "through the eyes of another",⁵⁴

It is important to the logic of this situation that the act of self-reflection by virtue of which the patient can emancipate himself from his illness, requires an encounter with an *other* through whom the patient can recognize the split off part of himself as his own. The patient's achievement of self-consciousness - which is to say the undoing of the unconscious which has become lost to the self - is *constitutively* an achievement of mutual recognition made possible by communication in a shared (public) language. Habermas gives this point its most convincing formulation in the following passage which appears as a footnote to the main text of *Knowledge and Human Interests*;

When the physician lets the patient free himself as an autonomous ego, the subjects must define themselves in relation to one another in such a way that the former patient knows that the identity of his ego is only possible through the identity of another who recognizes him and whose identity in turn is dependent on his recognition.⁵⁵

For successful therapy to be possible - for self-reflection to have its emancipating effect -

it must have a real affective-motivational base in the analysand. At one level, there must be a passion for self-knowledge which is strong enough to overcome resistance. But more fundamentally, Habermas remarks that there must be a "passion for critique",⁵⁶ in the sense that in order to emancipate himself from his illness through self-reflection, the patient must take *responsibility* for that which, like Hegel's criminal, "was to be in him but which is not". And it is this adoption of "moral responsibility for the content of the illness" demanded of the patient which results, by way of that form of self-reflection called scenic understanding, in the shattering or disenchantment of the experienced causality of fate;

For the insight to which analysis is to lead is indeed only this: that the ego of the patient recognize itself in its other, represented by its illness, as in *its own* alienated self and identify with it. As in Hegel's dialectic of the moral life, the criminal recognizes in his victim his own annihilated essence; in this self-reflection the abstractly divorced parties recognize the destroyed moral totality as their common basis *and thereby* return to it.⁵⁷

I now want to take up two sets of objections against Habermas's position. As we have seen, Habermas takes the metatheoretical province of psychoanalysis to lie beyond that of the cognitive/instrumental sciences. This is because psychoanalytical explanations are applied to the domain of human, symbolically mediated self-formative processes. The first set of objections takes issue with this move by challenging the appropriateness of Habermas's Hegelianizing of Freud on naturalistic grounds. The second type of objection, on the other hand, questions the compatibility of Hegel's insights with Habermas's analysis of systematically distorted communication as transcending the province of merely hermeneutic reflection.

(3.3) Naturalistic Objections

In his book *The Politics of Social Theory*, Russell Keat criticizes Habermas's conceptualization of the presuppositions of psychoanalysis for being insufficiently naturalistic in its construal of; (1) the domain of therapeutic understanding, and (2) the aim or goal of therapy.⁵⁸ By taking the domain of therapeutic understanding as the realm of a causality of fate, Keat claims, Habermas posits an incoherent alternative to the world of natural, causally determinate laws to which the human being, as a part of nature, is subject. This in turn blinds Habermas, Keat suggests, to therapeutic practices which rely for their success on technologically exploitable knowledge of nature. Keat then imputes this blindness to a misleading conception of emancipation which Habermas takes to be the goal of analytical therapy and, by implication, of Habermas's model of Critical Theory generally. I will now briefly consider each of these objections in turn, focussing as narrowly as possible on the weaknesses in Habermas's theory attributable to it, according to Keat, on account of its employment of the idea of a non-naturalistic causality of fate.

(1) Keat interprets the object domain of Habermas's critique as "the apparent objectivity of alienated, reified human *subjectivity*".⁵⁹ In his reading of Freud, Habermas identifies this split-off, alienated subjectivity with the unconscious as such. Habermas maintains, as Keat puts it, that "in overcoming the power of the unconscious, the patient is thereby removed from the deterministic realm of causality".⁶⁰ The imputation of this view to Habermas is supported by the following passage:

In technical control over nature we get nature to work for us through our knowledge of causal connections. Analytic insight, however, affects the causality of the unconscious as such. Psychoanalytical therapy is not based, like somatic medicine, which is 'causal' in the narrower sense, on making

use of known causal connections. Rather it owes its efficacy to overcoming causal connections themselves.⁶¹

As Keat interprets it, this commits Habermas to the view that a critical science (for which psychoanalysis is the model) is directed towards the "abolition of what only appear to be genuinely causal determinants".⁶² Against this view, Keat objects first that there are many manifestations of the unconscious which cannot in any plausible way be construed as 'pathological' - something to be 'abolished' - and hence as part of the domain of therapeutic reason. Dreams, jokes, and parapraxes, for instance, hardly seem apt targets for abolition. Second, he objects to Habermas's identification of alienated, unfree subjectivity with causally determined action. The causal determinations of human behaviour are, in principle, as susceptible to scientific investigation as any other part of nature. But this position does not entail, *pace* Habermas, the abolition of human freedom, since knowledge of such causes can remove the hindrances to purposive action. The domain of therapeutic reason is not pseudo-causality, but the competing causalities of different natural desires, instincts, and motivations.

Apropos of the first objection, Habermas explicitly accepts that the dream is "the 'normal model' of pathological conditions".⁶³ His position is that dreams follow a pattern which has the same structure of "pathologically distorted meaning" as symptoms - namely, the disguised manifestation of motives which have been excluded from public communication (i.e. 'repressed'). So although dreams are the 'normal' case, they are nevertheless only possible given an infantile history of conflictual and traumatic will-formation.⁶⁴ If the pathological state of affairs is taken as the *original* splitting off of need-disposition from communication, then it makes sense to say that dreams are pathological *in origin*. But this

is not to say that the dream itself is a pathology; on the contrary it fulfils the necessary compensatory function of repressed wish-fulfilment. Although, of course, something of the same is true of symptoms, in their case the compensatory mechanism issues in a *crisis of identity* which impels a certain 'passion for critique'. The point of the therapy is not simply *to remove* the symptomatic behaviour. It is rather to enable a *reconciliation* of the conscious agent with his unconscious or 'split-off' past. If Habermas's account is to survive the objection, then, it must distinguish between the pathological context or origin of the formation of the unconscious, and the pathological behavioural consequences of its formation. Of course, this presupposes that the origin *can* be determined, a supposition which will be questioned in the next section.

The force of Keat's second objection turns partly on the interpretation he gives to the passage he cites from Habermas, and specifically the meaning of the operative "overcoming" of causal connections. This overcoming, it must be made clear, is not achieved through the Kantian supposition that the category of causality is only applicable to the realm of phenomena - as distinct from the noumenal realm of the transcendental will. But only if Habermas were claiming this, would Keat's objection have force. What Habermas *does* claim is that since the human self-formative process is mediated by language, peculiar constraints are imposed on the categories which can legitimately be used for comprehending its disturbance. The crucial point is that at the therapeutic level, 'overcoming' the disturbance renders articulate what was otherwise unfathomable, and this in a process of dialogical interaction which both heightens self-consciousness and motivates altered behaviour through the affective content of what is brought to consciousness. The overcoming requires some responsibility on the part of the subject, who must act as if

things could be otherwise if the therapy is to be efficacious. This is different to a patient acting as if a cancer tumour is not going to develop, and it is a difference which needs to be reflected in the meta-theoretical self-understanding of the critical science.

Undoubtedly, Habermas's equivocal use of "causal connections" in the passage cited above is unfortunate. But I don't see why Habermas's position taken as a whole is incompatible with Keat's insistence that the therapeutic achievement lies in the replacement of one cause of action for another. All that Habermas requires at this point in the argument is that there are some motivations for action which, once rendered perspicuous by self-reflection, would lose their motivational power upon the subject's action.

(2) The second of Keat's objections I want to note concerns Habermas's specification of the goal of psychoanalytical therapy. Although this objection isn't particularly 'naturalistic', it does, I think, betray a fundamental miscomprehension of the role of 'the causality of fate' in Habermas's theory which is otherwise the target of his naturalistic criticisms. There is a normative complexity to the therapeutic process, Keat claims, which is inadequately captured by Habermas's notion of emancipation. Keat illustrates this objection with two hypothetical cases: of a male patient who is 'emancipated' from his emotional distress in his discovery that he feels insufficiently nurtured in his relationships with women - and then proceeds to impose stereotypical, oppressive roles on his wife; and of a patient whose 'successful' therapy reinforces a self-assertive attitude, which then motivates, and implicitly justifies, him in individualist and exploitative practices in his work-place. These examples show, states Keat, that "achieving therapeutic autonomy is consistent with the adoption and practice by an 'emancipated' patient of attitudes and values that are by no means [from an

anti-sexist, socialist viewpoint] unobjectionable". "And I do not see", Keat continues, "how their acceptability or unacceptability can be determined without going well beyond the normative concepts illustrated by the model of psychoanalysis".⁶⁵

But emancipation as it is understood according to Habermas's Hegelian reading of Freud is precisely not normatively naive in the sense Keat alleges. For it involves recognition of the individual's immanence within a virtual ethical totality which is *itself* avenged in the individual's self-alienation prior to the therapy. Emancipation from unconscious dependencies represents only *one* moment of the normative structure which Habermas takes to be illustrated in the model of psychoanalysis. As Habermas puts it, "the virtual totality that is sundered by splitting-off is represented by the model of pure communicative action",⁶⁶ but alienated subjectivity is only one consequence of this rupture. Thus the key normative concept which Habermas takes to be illustrated by the model of psychoanalysis is intact or undamaged *inter*-subjectivity, as represented by the concept of pure communicative action. The exploitative practices which Keat suggests are consistent with the achievement of therapeutic success *themselves* transgress the normative model of an intact intersubjectivity to which Habermas appeals; as the analogy with the criminal who revokes the principle of ethical community makes clear. There are *other* crises in the symbolically mediated process of the identity-formation of humans, claims Habermas, which result from systematically distorted communication. But clearly there is also an asymmetry here. For what is apparently missing in the cases cited by Keat, is an account of the *motivational* basis for recognizing the claims of others in non-exploitative action. This leads us to different kind of objection to the understanding of human identity-crises Habermas draws from psychoanalysis.

(3.4) Anti-Naturalistic Objections

In the previous section, I considered an argument to the effect that the weakness in Habermas's appropriation of psychoanalysis as a model of therapeutic reason lies in his failure to address the theoretical implications of the fact that psychopathologies have natural causes. But Habermas's depth hermeneutical reading of Freud can also be criticized for being too naturalistic in its conception. This is because for Habermas, psychoanalytical self-reflection acquires its explanatory power by virtue of combining hermeneutical interpretation of apparently incomprehensible behaviour with empirical scientific insight into the causal origin of that incomprehensibility. To be sure, the imputed explanatory scientific insight does not refer to an event covered by a law-like generalization, but to disturbances in the early stages of a linguistically mediated self-formative process. Since these disturbances are themselves linguistic, they are curable by hermeneutic, self-reflective means, rather than by the technical manipulation of the efficient causes determining the behaviour of natural objects.

Nevertheless, Habermas supposes that *there is* a determinable split-off symbol and repressed motive which avenges itself in the fate of the patient, and that this fate is 'disenchanted', to emancipatory effect, when it is 'recovered' in the analysis. Only on the supposition of a re-internalization of given excommunicated needs, Habermas believes, can the continuation of a disturbed self-formative process count as the criterion of validation for the theoretically constituted narrative of self-formation. Following arguments recently put forward by Jay Bernstein, I will now consider (1) the plausibility of the naturalistic assumption built into the former thesis, and (2) its compatibility with the hermeneutic claim expressed in the latter.

(1) Borrowing a formulation from Amelie Rorty, Bernstein describes the task of psychoanalysis, with which Habermas agrees, as that of identifying "the intentional component of the significant cause of the dispositional set that forms the intentional component of the [anomalous, intractable, misery-inducing] emotion".⁶⁷ The intentional component of the patient's neurotic feelings is not their apparent object, and this is something of which the patient is often aware. Rather, the object of the emotion refers back to a disturbance in the dispositional state of the patient, the particular nature of which the patient is not aware. This is because the disturbance itself is taken to have a remote cause. But this is a significant cause in virtue of its intentional component; there are objects to the traumatic emotional conflict of the original childhood scene. As a result of this conflict, Habermas holds that the intentional object of the emotion becomes split-off from the emotion. The disturbance in the dispositional set which results from this renders the actual meaning of the neurotic symptoms intractable. In the transference scene, according to Habermas, equivalence is established between the everyday expression of the disturbed need disposition in symptoms, and the original significant cause of that disturbance.

But as Bernstein notes, this claim to equivalence fits ill with the vague, underdetermined nature of the intentional components of the misery-inducing emotions in question. The specification of the intentional component always falls under some interpretative schema; what counts as a correct description of it is internal to the particular interpretative theory employed by the therapist. This interpretative theory-ladenness applies both to the correct description of the latent intentional object of the currently disturbed need disposition, and to that of the remote significant cause of the disturbance. But in neither case is the

identification simply a matter of neutral, empirical discovery. Although at first sight the former may appear more amenable to theory-neutral description, the criteria for what counts as a *relevant* description will in turn be drawn from the latter, highly theoretically loaded, reservoir of interpretive terms. But this renders the claim to equivalence highly problematic, since in the case of the symbolic intentional components of intractable emotion and need disposition, what counts as equivalent is not determinable independently of the interpretative schema through which they are identified. Worse, they can hardly be said *to exist* independently of such a schema. Unlike empirically equivalent but incompatible theories, "what these reflective theories are about", Bernstein suggests, "becomes different as the theories are accepted and so become true".⁶⁸ But if there is a truth of the matter here, if psychoanalytical self-interpretations do admit of cognitive validity, and if the existence of the object-domain of psychoanalytic self-reflection "is contingent upon the acceptance of the theory", in what can this validity consist, and in what sense can that domain be said to exist at all?

(2) The validity, as we have seen, cannot reside in the re-representation of a split-off symbol; the conditions of acceptability of an analytical interpretation do not lie in an accurate description of a determinable past. Instead, Bernstein insists that the truth of the self-reflection is essentially a truth of a *practical*, productive character, discernible retrospectively in the act of self-narration. It is fundamental to Habermas's own view that the analytical self-reflection is undertaken in the context of some practical need; there must be a passion for critique - namely, to resolve the place in one's life of certain feelings, emotions, and irrational behaviours. For this goal to be realizable, the conditions of acceptability for analytic insights must have some affective basis. What is required is not

just correct beliefs about the past, but rather the significance of dispositions as they are revealed to play a role in one's life as a whole in the autobiography which the theory mediates. But this 'as a whole' projects into future life; it orients the agent *to become* the kind of person who, according to the interpretative schema of the theory, has reached moral maturity. What it is to reach moral maturity cannot be separated from the self-interpretative framework through which the *self-formative process* is grasped, through narrative form, with practical intent.

This point will be elaborated in greater depth later, when its place in Taylor's hermeneutic approach to the moral predicaments of modernity will be explored. As is clear from his reading of Freud, Habermas does himself affirm the validity of narratively carried-out self-critique, but by presenting it as attempting to dispel particular illusions within the totality of a course of life by way of a scientific explanation of their causal origin, he opens himself to the charge that the meaning of the totality of the analysand's life is *itself* changed by the autobiographical practice through which those practically significant illusions are identified. This is because the practical task of dispelling those particular self-deceptions is 'always already' oriented to the horizon of selfhood immanent to the interpretative framework of the theoretically mediated self-reflection: the very identification of the cause of the disturbance in the self-formative process is internal to the particular theory which mediates the autobiography. To return to the third and fourth features of the concept of crisis introduced in (1.2), the criterion of emancipation by appeal to which crises are normatively ascribed, and the emancipatory effect of crisis-resolution, is not independent of the framework of interpretation within which the self regains its identify.⁶⁹

(3.5) Conclusion

What then is it for the causality of fate to be disenchanting? We can now distinguish two apparently competing answers offered by Habermas. Both appeal to the idea that the causality of fate appears 'behind the backs' of subjects, and requires *theoretical* insight in order to be disenchanting. To the extent to which Habermas draws from the young Hegel, he claims that this insight into the common interconnection of lives - the shared basis of existence - when supported by an effective 'longing' for what has been lost when this basis is revoked, *thereby* effects reconciliation with the ruptured ethical totality. This form of disenchantment through insight we might call 'love's knowledge'. On the other hand, however, Habermas appeals to a form of theoretical insight which disenchanting by virtue of exposing pseudo-objectifications - by dispelling the illusion of nature-like necessity from which emancipation can be gained through insight into the linguistic mediation of subjects. This kind of insight reconstructs distortions in the linguistic mediation of subjects, and the crises which issue from them, within the framework of a theory of communicative competence. This is then construed by Habermas as a theory of the general capacity for communicatively rational action, from the perspective of which pseudo-objectifications can be theorized as the prevention of the procedural realization of rational life together by the functionally rational steering media of money and power. But on the way to accounting for the disenchantment of the causality of fate within the problematic of the paradox of rationalization, Habermas has adopted a conceptual apparatus which in turn problematizes the intelligibility of the kind of insight originally derived from Hegel. We seem to have competing conceptions of the self-clarification of the claim to normativity.

My aim up to this point has been to clarify the form and dynamics of crisis formation

which the causality of fate model of critique is employed to illuminate. As we saw in chapter one, Habermas distinguishes three different contexts of application for the concept of crisis; the medical, the tragic, and the social scientific. In each case, Habermas claims that the ascription of a crisis has some normative force: the diagnosis of a critical state presupposes some appeal to a model of healthy or mature functioning, and the successful resolution of a crisis issues in a form of emancipation. Habermas's social scientific concept of crisis organizes a theory which attempts to reconstruct the self-formative process of the human species, a process which, when disturbed, issues in socially induced crises of identity. A moral order, claims Habermas, is immanent to the process of human self-formation; and this is shown in the identity crises which befall humans when the moral order is transgressed. This ethical totality, we saw, is represented by the model of 'pure communicative action'. In the next chapter, I will consider this concept in more detail. I will then argue that the key difficulties noted in this chapter arise from his attempt to render scientific the tragic conception of crisis upon which he ultimately draws in his attempt to solve the problem of self-reassurance. But before that, I must address the conception of language on the basis of which Habermas constructs a standard for critique as the rational foundation of the claim to normativity.

CHAPTER FOUR: HABERMAS, LANGUAGE AND CRITIQUE

(4.0) Introduction

In (2.4), I noted that Habermas's strategy for resolving the modern tension between cognition and identity turns on establishing two claims. The second was that the norm of 'intact intersubjectivity', which Habermas proposes as the source of modernity's critical self-reassurance, can be traced back to the vehicle for reproducing modernity's symbolic resources; namely, in the procedural conditions of the communicative use of propositionally differentiated language. This is related to the first claim I identified, that modern conditions of social life are disposed to crisis tendencies and pathologies, in the following way. Human identities can only be acquired and maintained within the context of a lifeworld. Following Habermas, the lifeworld is the source of those cultural traditions, group solidarities, and individuating communicative competences, which condition the possibility of human identity. But Habermas asserts that the reproduction of these three basic structures of the lifeworld "*can take place only through the medium of action oriented to reaching understanding*".¹ Were the lifeworld to be reproduced non-communicatively - either by what Habermas calls the 'strategic action' of individuals who are oriented to ego-centric success according to criteria of instrumental rationality, or via the systemic media of money and power, operating according to criteria of functional rationality - then those sources of identity which constitute the lifeworld would not be sustainable. And the 'inevitability' of such a disintegration of the lifeworld is precisely that of a 'fate' which is a matter neither of objective law-like necessity nor of subjective decision;

That communicative rationality, precisely as suppressed, is already embodied in the existing forms of interaction and does not first have to be postulated as something that ought to be is shown by the causality of fate which Hegel and Marx, each in his own way, illustrated in connection with phenomena of ruptured morality - the reactions of those who are put to flight or roused to resistance by fateful conflicts, who are driven to sickness, to suicide, to crime, or to rebellion and revolutionary struggle".²

In the previous chapter, I considered how Habermas conceptualizes mental sickness and crime as social or intersubjective pathologies following a logic of fateful causality and conflict. But what are we to make of the status of the claim that communicative rationality "is already embodied in existing forms of interaction" as an 'is' and not just as an 'ought'? Does Habermas have a philosophical argument to justify the claim that communicative rationality is 'always already' embodied in modern forms of interaction, and that such interaction can generate *out of itself* normative standards of rational critique? If so, and if the argument is a good one, Habermas might be said to have solved the philosophical problem of the self-reassurance of modernity as he has defined it.

There is an argument, and the key to it lies in the distinction between, and the relative priority of, two fundamental types of linguistic interaction. As beings whose identity is mediated by language, in cultural traditions, group solidarities, and communicative competences, linguistic interaction is the *sine qua non* of any identity for that kind of being at all. I will begin then (4.1), by drawing out the distinction Habermas makes between these two different kinds of linguistic interaction; what he calls "communicative" and "strategic" action. Now not only do these two kinds of action have *competing* criteria of rationality, but communicative rationality, we have just seen, is *suppressed* in existing modern forms of interaction - forms that presumably accord with criteria of strategic

rationality.³ How is this to be explained? Habermas replies by way of arguing that communicative interaction is in some sense "originary" or intrinsic to language use; that the linguistic coordination and mediation of subjects and their actions would not be possible if communicative rather than strategic language use were not the 'original' mode of linguistic interaction. I consider this part of the argument in (4.2). This is preceded by some remarks justifying the interpretation of Habermas I offer, which by highlighting the importance within Habermas's theory of establishing the 'originary' status of communicative action, diverges quite sharply from some current readings of Habermas's work. Its significance only comes to light, I propose, when it is grasped in relation to the theme of the 'causality of fate' which guided the previous chapter.

I then examine two sets of objections to this thesis which are particularly relevant to the broader problem of modernity's self-reassurance (4.3); what I call the 'deconstructivist objection' and the 'agonistic objection'. While I argue that Habermas's position is not significantly threatened by these objections, his thesis concerning the primacy of communicative action is seen to be in need of further support. In (4.4) I offer a reconstruction of an alternative argument implicit in Habermas's text which might provide it. This alternative argument relies on insights drawn from Wittgenstein's famous remarks on rule-following. Identical ascriptions of meaning, it is supposed, are a condition of possible successful linguistic communication. But identical ascriptions of meaning are only intelligible, Habermas argues, if certain idealizations are built into the pragmatics of language use. In (4.5) I question the coherence of Habermas's use of Wittgenstein's rule-following considerations, by introducing a set of 'Wittgensteinian objections'. I conclude by suggesting that they pose a serious threat to the acceptability of Habermas's

thesis concerning the primacy of communicative action, and the role it plays in his contribution to the philosophical discourse of modernity.

(4.1) Communicative and Strategic Linguistic Interaction

Habermas argues that the norm of an intact intersubjectivity can be derived from the structural presuppositions of what he calls 'communicative action'. Despite the immense theoretical burden carried by the concept of communicative action, as no less than the organizing idea of Habermas's critical diagnosis of modernity and its ground for self-reassurance, the concept suffers from a widely noticed lack of clarity and consistent application.⁴ This concept is not amenable to simple definition, but one can most readily grasp its point if one considers the basic need for human beings to coordinate their actions.⁵ If any but the most rudimentary actions are to be coordinated, they must be so via the medium of language. And in the process of linguistic interaction - in the medium of which actions are coordinated in a society - human beings are socialized and take on individual and collective identities. So human identities are mediated intersubjectively by language. The point of the concept of communicative action is both to explain this process, and to diagnose systematically its tendencies to go wrong. By 'go wrong', I mean a failure of the mediating mechanism to integrate and sustain identities. Such tendencies thus constitute a disposition towards crisis.

Insofar as competent linguistic interaction is a condition of having an identity at all, then the reconstruction of the norm of an intact intersubjectivity from the conditions of identity formation requires a distinction between those kinds of interaction which can, and those which in principle cannot, sustain the identities of human beings *qua mediated*. To this

effect, Habermas contrasts communicative action with strategic action. The arguments offered in support of this demarcation are controversial, but their force is missed, so I argue, if they are taken independently of the model of 'dirempted' intersubjectivity considered in the previous chapter. I shall proceed by briefly outlining the general distinction between communicative and strategic action, before turning to the more specific, and for our purposes more significant, distinction and relation between communicative and strategic linguistic interaction.

Habermas proposes that communicative action can be contrasted to strategic action by virtue of its peculiar 'attitude' or 'orientation':

social actions can be distinguished according to whether the participants adopt either a success-oriented attitude or one oriented to reaching understanding. And, under suitable conditions, these attitudes should be identifiable on the basis of intuitive knowledge of the participants themselves.⁶

(Strategic) action oriented to 'success' and (communicative) action oriented to 'reaching understanding' are distinct types of action, rather than merely different aspects of the same action. These distinct action orientations are *in principle* identifiable by the person who is acting, but as the 'under suitable conditions' proviso makes clear, this *need* not be so. For instance, the background or pre-theoretical knowledge of the participants must be sufficiently well developed and differentiated to enable them to distinguish between strategic 'success' and communicative 'reaching understanding'.⁷ Now there is a problem with resting this distinction upon an agent's intention which will only become clear by the end of the chapter. Let us begin by taking a closer look at what is being distinguished here.

When an actor engages in strategic action, he seeks to influence an opponent in pursuit of an end he himself has decided. Strategic action is therefore a species of instrumental action, where the means employed to realize a given end involves another person. The orientation of strategic action is success in bringing about a desired outcome by means of another actor. The only thing which is constitutive of the point of strategic action is its utility or consequence for the strategic actor. The rationality of this kind of action is determined by the degree of success the actor achieves (or could be expected to achieve, given the relevant information available) in realizing his intended outcome. According to this conceptualization then, the telos of strategic action is to maximize the actor's own self-interest or "egocentric utility".⁸ Insofar as the action-orientation is towards the successful implementation of an individually decided plan, and another actor enters into the meaning of the action solely by virtue of being a means to that monologically (rather than through dialogue) decided end, the action is strategic in type. And to the extent that an actor achieves success by "causally exerting an influence upon others"⁹, by an influence which empirically forces them to act in accord with the strategic actor's goal, the action can be described as strategic. The coordination between strategic actions is typically mediated by threats and rewards. The interaction between the torturer and his victim, the pimp and the prostitute, and the prostitute and her client, are extreme but graphic examples of this kind of action coordination.¹⁰

Habermas refers to communicative action, on the other hand, as action with the attitude or orientation of "reaching understanding" (*verständigung*). By "reaching understanding", Habermas means agreement reached between at least two actors on the basis of its mutually acceptable validity. As an agreement, it must have the assent of both parties. The meaning

of a communicative action is thus constitutively dialogical, since such an action is specifiable only under the conditions of an agreement made *with* another. It is also dialogical in the sense that the agreement is reached solely by virtue of that force which is peculiar to dialogue - the force of the better argument or the 'pull to validity'. The acceptance of what is 'offered' in a communicative action is rationally based, precisely, because it is based upon the acceptability of what is offered. What is offered in communicative action is the content of an utterance. According to Habermas's model, when I engage in communicative action with another, I make a speech-act which raises a validity-claim. A speech-act is something I do with a sentence or a proposition. In communicative action, I either implicitly or explicitly offer a proposition as something which has a claim to validity. My partner in communicative action can reject or accept the validity claim I raise - s/he can take a 'yes' or 'no' position on it. This position is not a mere *de facto* acceptance (or rejection), but one based, Habermas insists, on the acceptability or validity of the claim. In either case, this position is based (again either implicitly or explicitly) on grounds or reasons. Since "validity claims are *internally* connected with reasons and grounds"¹¹, they are *criticizable* with respect to whether the validity conditions of the claim are satisfied. And the claim is valid or invalid depending upon whether the validity conditions of it can be shown to be fulfilled (or 'redeemed') through what Habermas calls "discourse" or "argumentation".

Since the agreement or consensus toward which communicative action is oriented concerns the content of a validity claim, the claim must be 'propositionally differentiated'. That is, the content of the claim must be differentiated out from its attitude. Communicative action is thus motivated by the attitude of reaching an agreement over the content of a validity

claim. A communicatively reached agreement, characterized by the internal relation between validity and reasons, must be distinguished from agreements 'causally induced by outside influence'. Now it is just such outside influences (or inducements) which strategic actions exploit - in, for instance, monetary rewards or threats of violence. It is thus of some importance that Habermas establish that the 'propositional attitude' - or more precisely, speech act orientation - of communicative action is *not* a 'causally induced outside influence'. Or if one insists that some kind of causality must be in play here, then one could say that the agreement or consensus to which communicative action is oriented obeys a causality of *reasons*.

Habermas is not arguing that, as a matter of fact, communication actually takes place according to this discursive procedure. The argument is rather counterfactual in form.

"Without doubt", he recognizes,

there are countless cases of indirect understanding, where one subject gives something to understand through signals, indirectly gets him to form a certain opinion or to adopt certain intentions...or where, on the basis of an already habitual communicative practice of everyday life, one subject inconspicuously harnesses another for his purposes, that is, induces him to behave in a desired way by manipulatively employing linguistic means and thereby instrumentalizes him for his own success.¹²

There are countless cases, then, of the strategic use of language in everyday life. Manipulation by linguistic means is a commonplace, as are matter of fact agreements or accords which facilitate action-coordination. But these agreements, Habermas insists, will not be based on *convictions* which stand the test of argument. They are not, that is to say, rationally motivated, and therefore not the kind of agreement by virtue of which actions are

coordinated communicatively. But Habermas wants to argue that these habitual practices of empirically motivated linguistic instrumentalization are in some sense derivative from the rationally motivated practice of communicative action, even though in any actual situation non-rational empirical motives and power relations are causally in play. Why should Habermas want to argue this?

The question can only properly be answered if it is raised in the context of the problem of modernity's self-reassurance. For if human beings are mediated by language, and language can be shown to be intrinsically or essentially coordinative as represented in the model of pure communicative action, then the damage to intersubjectivity in modernity attributable to the mediating mechanism can be criticized by appeal to this model. Pure communicative action, as we have seen, is the *virtual* totality which is sundered and which avenges itself in the dynamic of fateful causality in the formation of Freudian psychopathologies.¹³ It is therefore essential that Habermas should establish that pure communicative action is *always already* (or 'virtually') at play in all linguistic interaction. He must show, that is, that communicative action is the 'original' mode of language use, and that strategic actions are 'parasitic' upon it.

This *must* be the case if the causality of fate is to be conceptualized as coming into play with the violation of the structure of rational life together, and if the fundamental structure of rational life together is to be conceptualized within a theory of the rationalization of action-coordinative mechanisms. For only then can the bare linguistic mediation of human beings assume its place in the self-reassurance of a rational (qua communicatively rationalized) form of life, the form of life which for Habermas is the goal of the project of

modernity.

It is therefore not enough, as Stephen White has recently suggested, that "Habermas should just admit that both [communicative and strategic] forms of interaction coordination are necessary to social life".¹⁴ If one fails to see that intact intersubjectivity requires communicative action, and that non-communicative but in some distorted sense 'rational' intersubjectivity generates the problem of self-reassurance, then like White one will "not see why this distinction [between 'original' and 'parasitic' language use] is so crucial".¹⁵ I will maintain that the distinction *remains* crucial insofar as Habermas's contribution to the philosophical discourse of modernity, the insight into and disenchantment of the causality of fate, is presented as part of an acultural theory of rationalization. Although another leading Habermas commentator and critic, David Rasmussen, does see its importance,¹⁶ he too fails to highlight where this importance really lies; precisely in spelling out the *conceptual priority* of language considered as a dialogically structured, mediating ethical totality.

As a sympathetic critic of Habermas, White is prepared to play down the significance of the thesis that communicative action is the original use of language because he accepts that it has suffered fatal criticism. I will now argue that this criticism takes the bait of a misconstrual of Habermas's own position. This position is only at all plausible if seen in the context of the causality of fate model of intact and distorted intersubjectivity, a model which, taken together with a theory of rationalization, appeals to a *causality of reasons* distorted by 'external influence'.

(4.2) The Primacy of Communication; A Bogus Argument

Habermas seeks to establish that there is a norm, captured in the model of pure communicative action in a dialogue situation, which is immanent or virtual to linguistic interaction and which can be appealed to as a source of critical self-reassurance for a form of life which has a well-grounded claim to rationality. To support this view, he argues that use of language which either presupposes or establishes an instrumental/ coercive relationship is 'parasitic' upon the procedure of pure communicative action. This he does by arguing that the strategic use of language, the use of language oriented to the successful manipulation and control of an opponent according to an individually decided plan, is derivative from the 'original' communicative use of language (oriented to reaching consensus). But at the point in which he expressly makes this claim, as several critics have noted, the argument offered is both confused and thoroughly unconvincing.¹⁷ I will use this section to indicate why.

If it is acknowledged that language can be used with a strategic-instrumental orientation to control and manipulate an opponent, the thesis that an intact intersubjectivity is virtual to language requires "it can be shown that", as Habermas regrettably put it;

the use of language with an orientation to reaching understanding is the *original mode* of language use, upon which indirect understanding, giving something to understand or letting something be understood, and the instrumental use of language in general, are parasitic. In my view, Austin's distinction between illocutions and perlocutions accomplishes just that.¹⁸

As Habermas presents it, an illocutionary act refers to what is done in delivering an utterance. Under 'standard conditions', that is, where everyday life contexts and conventions are not bracketed (as, say, in a play), its meaning can be traced to a (literally

significant) performative. Such performatives include assertions, commands, promises, and avowals. *In* saying, "French philosophers are charlatans", "I will pay you back tomorrow", or "I love you very much", under standard conditions I make the illocutionary acts of performing a declaration of belief, a promise of action, and *inter alia*, an expression of feeling. The performative verb of the illocution, what is done in saying it, can be prefaced by "I hereby...". In these cases, "I hereby..." declare, promise, avow.

By contrast, the perlocutionary act refers to what is done *through* or *by* saying something, to the intended effect brought about in the hearer as a consequence of the utterance. By saying those things, for instance, I might be trying to wind up a stylish literary critic, to get my bank manager to leave me alone, or to get my lover not to leave me alone.¹⁹ But unlike the illocutions, only *I* will know if these are faithful descriptions of my respective perlocutions. The significance Habermas draws from this is as follows.

Illocutions are said to differ from perlocutions in their respective aims and the expressibility of those aims. The aim of the illocution is manifest in the meaning of the performative verb. "I confess" makes manifest my confession. In going no further than the manifest meaning of what is said, insofar as this meaning is regulated by public linguistic conventions, the aim of the illocutionary act is 'self-identifying'. The addressee does not need to go 'behind' the speech-act in order to find the meaning of it as the communication of a greeting, a confession, etc. But if the illocutionary act is constituted by the meaning of what is said irrespective of the agent's subjective intentions, the aim of the perlocutionary act is not self-evident in what is said, since it depends on what consequences the speaker intends to bring about through the act. The illocutionary aim of "please get me a glass of

water" is self-identifying in the making of a request, though the perlocutionary aim, what I want to get *through* making the speech-act, is my own intention to which I have privileged access. I may be thirsty, I may be buying time, I may be wanting to get rid of the addressee, and so forth.

Since the aim of the illocutionary act is self-identifying, the conditions of its success can be read off from a description of the act which makes the performative explicit. The case is different with a perlocutionary act, Habermas asserts, since the aim goes beyond linguistic conventions to the 'teleological context' of the agent's strategic intentions. This asymmetry, Habermas argues, is shown in the fact that the 'success' of the illocutionary act is not compromised by the illocution being openly declared (as a promise, a command, a confession etc.), whereas an express declaration of what would count as perlocutionary success may result in the failure of the perlocution. Habermas redefines perlocutionary acts as a subspecies of teleological (means-ends) action where the means is a speech act and the end is undeclared.²⁰

But can one conclude from this that strategic linguistic discourse is somehow derivative from communicative action? Habermas seems to be saying here that if perlocutions (speech acts with perlocutionary force) are derivative from illocutions (speech acts with illocutionary force), then *eo ipso* strategic speech-action is shown to be parasitic upon communicative action. But this move would only work if the distinction between communicative and strategic action were either synonymous with or strictly analogous to the distinction between illocutionary and perlocutionary acts. And only if the distinction were of this kind would it be legitimate to make the inference Habermas seeks. So even if the conditions of

perlocutionary success can be shown to presuppose a successful illocution, this as yet tells us nothing about the relationship between communicative and strategic action, since the distinctions are not synonymous, and a relation of strict analogy has not been shown. And if the former distinction is transformed in such a way that it does become synonymous with the latter, then the question is begged, since the putative argument is to *justify* the claim that communicative action is the original mode of language use by appeal to Austin's distinction.²¹

The actual role of the distinction between illocutions and perlocutions must be interpreted as an *elaboration* of the distinction between communicative and strategic action, rather than as a *justification* of the priority of the former over the latter. Consider where we left off the previous section. I said there that the 'causality' of linguistically mediated interaction is one of reasons, and that this medium is distorted by influences 'external' to it. These rather opaque remarks should now be clearer, for their meaning is bound up with the distinction between the illocutionary and perlocutionary force of speech acts. The illocutionary force of a speech act is internally related to the meaning of the linguistic utterance, and achieves its coordinating effect between actors by virtue of this connection. Since, for Habermas, a speech-actor understands the meaning of the claim raised in an utterance when s/he grasps the conditions of its validity, the causality of the illocutionary effect is validity- or reason-conditioned. The perlocutionary effect, however, owes its existence not to the intrinsic meaning of the speech act offer, and hence to its validity basis, but to causal powers independent of the (often implicit) warranty to justify the content of the utterance.

The virtual totality of the dialogue situation then, refers to linguistic interaction in pursuit of exclusively illocutionary effects. Whereas perlocutionary effects "are intended under the description of states of affairs brought about through intervention in the world", illocutionary results appear "in the lifeworld to which the participants belong They cannot be intended under the description of causally induced effects".²² Rather, the illocutionary force of the speech act lies in the bond resulting from the warranty to justify the validity claim offered. And this is conceptually independent of the empirically induced bonding of perlocutionary effects. The perlocutionary force of a speech act stands to the content of what is said as a *communicatively* 'irrational force' in that it is exerted independent of the redeemability of the validity of that content. But if the distinction between illocutions and perlocutions has an elaborative rather than a grounding function in the argument for the primacy of communication, is even that function well served by the distinction as Habermas has drawn it? It is far from clear that a substantive demarcation between illocutions and perlocutions can be established according to the criteria Habermas suggests. It does not look very plausible, for a start, to argue that perlocutionary aims *must* be undeclared as a condition of their success. There are certainly cases where admission of my plan may be self-defeating for its successful execution. If I were to say "I hereby promise to reduce base lending rates and by saying this I am trying to get you to vote for me", this may well jeopardize the success of my plan to get votes. But equally, the person I am trying to influence may easily recognize my strategic intent even if I don't make it explicit, so that it simply doesn't matter if my plan is declared or not.²³ Further, I might even enhance the chances of success in my perlocutionary goal by openly declaring it. For instance, if my goal is to remove someone from my company, I might say to that person "I hereby declare that the sight of you fills me with disgust", but my chance of success may

well be enhanced if I were to add "and by saying this I am trying to get rid of you".²⁴

The general point to be made here is that it depends on the circumstances whether expression of the perlocutionary aim defeats that aim. But although I think Habermas could drop (or modify) the intention-concealment requirement of perlocutionary success without having to drop the illocutionary/perlocutionary distinction altogether, the general objection that this distinction is context-specific brings me to a more serious charge concerning the *kind* of context Habermas appeals to when illustrating perlocutions.

Consider the following passage where Habermas argues that the difference between illocutionary aims, the success of which is conditioned by open expression, and perlocutionary aims can be seen in the fact that;

the predicates with which perlocutionary acts are described (to give a fright to, to cause to be upset, to plunge into doubt, to annoy, mislead, offend, infuriate, humiliate and so forth) cannot appear among those predicates used to carry out the illocutionary acts by means of which the corresponding perlocutionary effects can be produced.²⁵

If such predicates as these were genuinely representative of perlocutionary effects, then indeed they would be best kept hidden for strategic purposes. But a much different picture emerges if rather than these antagonistic, *non-cooperative* perlocutionary effects, others such as 'to give relief to', 'to cause to be uplifted', 'to reassure', 'to boost confidence', 'to guide' were to be taken as exemplary. That they are not suggests that Habermas elides over such examples because he is *already* committed to the cooperative function of illocutions. And since he wants them distinct from perlocutions, he is blinded to the latter's cooperative potential, and hence fails to consider such instances of it.

My point here is not the one made by Culler that a speaker can also seek (and achieve) strategic success with illocutions such as orders and commands.²⁶ According to that objection, the speech-act "I hereby command you (Y) to do x" will be successful if Y consequently *does* x, and thus there is no difference between the criteria of illocutionary and perlocutionary success. For Habermas can reply here that in the case of imperatives, where an agent openly declares his illocutionary aim to command someone, influence is exerted without reference to a raised and criticizable validity claim. The addressee is not in a position to rationally accept or reject the offer of the speech act. To be sure, the criticism is telling if Habermas is taken at his word in justifying the priority of communicative over strategic speech acts by reference to the illocutionary/perlocutionary distinction, since orders are, after all, illocutions. But it should be just as clear that orders which do not raise criticizable validity claims cannot count as communicative acts, even though they have illocutionary force. As I have already suggested then, Habermas should not be taken at his word. He is in fact redefining perlocutionary acts as concealed strategic acts governed by criteria of instrumental rationality, as opposed to genuine communicative acts oriented to reaching understanding. But perlocutions need not be as antagonistic to reaching understanding as Habermas's illustrations of strategic action suggest. To illustrate my preceding point, a teacher may use perlocutionary acts to encourage self-confidence in students, with the (possibly concealed) aim of enabling the pupil to reach an understanding.²⁷

Habermas might reply to this that I am not using the term 'reaching understanding' in its proper technical sense. For I am neglecting the condition that a communicatively reached agreement is based on mutually acceptable reasons. And it might be said, that the

distinction between illocutions and perlocutions really boils down to this. Illocutions exploit the validity (rational) dimension of language which is irreducible to the causal (strategic) dimension of language exploited by perlocutions. This thesis is untouched by my objections so far. But an argument used by Habermas to support this thesis has, I think, been undermined. This is the argument that irreducibility is shown in the essential antagonism between validity-based and empirically motivated linguistic interaction. But of course the absence of this antagonism does not entail that rationally motivated interaction thereby collapses into causal influence. It does, however, threaten Habermas's fundamental idea that forces 'external' to pure communicative action are responsible for distortions in the linguistic mediation of the subject.

It will be recalled that an objection of just this kind was seen to be applicable to Habermas's interpretation of Freudian psychopathologies according to Hegel's model of a causality of fate. This connection helps us to grasp the point of the imputed primacy of communication. For Habermas communicative action is the original mode of linguistic interaction for the same reason that, for Hegel, the ethical totality precedes the criminal act which revokes it. The transcendental argument is that *given* the existence of crisis experiences - of a rupture with an ethical totality or presupposed cooperative community - the coordinative operation of language is 'originary'. It is original in the sense that hidden conflict and manipulative strategy are 'external' to the 'binding force' of intersubjectivity which derives from the warranty to justify the content of speech-acts. The generation of modernity's norms out of itself thus means their generation *exclusively* from the domain of the lifeworld - from the force of illocutions made without reservation. But this is only possible given the uncoupling of the lifeworld from the causal inducements of the system, and is therefore only possible

in the condition of modernity. For the symbolic reproduction of modern societies depends upon identity-formation via communicative action, a thesis corroborated by the pathological consequences of non-communicatively mediated action coordination in the contemporary world. But self-reassurance can be attained if it can be shown that immanent to this uncoupling are identity-securing norms built into the possibility of social integration via communicative action. Habermas's theoretical strategy has been to show that this immanence resides in the intrinsically consensual nature of language.

(4.3) Consensus and Strategy in Language: Two Sets of Objections

I now want to take up two different kinds of criticism which have been launched against Habermas's approach to consensus and strategy in language. The first takes issue with the very idea of an 'original' mode of language use by way of highlighting the implicit bias involved in the distinction between 'normal' (consensus-oriented) and 'abnormal' (non-consensus oriented) language. We can call this the 'deconstructivist objection'.²⁸ The second kind of objection goes by way of offering an alternative model of language which emphasizes and prioritizes precisely its strategic aspect. This view, put forward by Lyotard, I shall call the 'agonistic objection'.²⁹ Although Lyotard draws upon certain Wittgensteinian themes, his position is quite alien to what can properly be called a Wittgensteinian one. But what can be so called is of no small importance to us. For as I shall go on to argue in the next section, an alternative support for the thesis of the primacy of communication can be teased out of Habermas's interpretation of what Wittgenstein establishes in his remarks on rule-following. I will then return to a third set of objections directed against the adequacy of this interpretation and its consequences for understanding the role of consensus in language. These Wittgensteinian objections can be extended to

Habermas's conception of critique, creating tensions within Habermas's proposal for a self-clarification of the modern tension between identity and cognition which will be examined in the next chapter.

(1) *The Deconstructivist Objection*: When Habermas states that communicative action is 'originary', he is claiming that it must be taken as the 'normal' case of linguistic interaction upon which other cases are parasitic or derivative. Habermas's analysis of communicative action appeals to a notion of a 'standard speech-act' which is uttered 'seriously' and used as 'simply' and 'literally' as possible in 'normal' everyday practice. Speech-acts not specifiable in this way - such as jokes, playful fantasies, imaginative role-playing, puns and ironies - are thus derivative in that they necessarily presuppose the already established communicative competence of 'reaching understanding' under 'standard conditions'. The deconstructivist objection challenges the assumption of such 'standard conditions'.³⁰ For the deconstructivist, there is no innocent realm of 'ordinary' language upon which other forms of discourse *can* be parasitic. What counts as 'standard' or 'normal' is little more than a reflection of a pre-decided evaluative preference of the theorist; in Habermas's case a particular conception of truth, rightness, and sincerity. Indeed the very term 'parasitic' shows the devaluation of the non-serious, the abnormal, and the marginalized in language. Habermas's distinction between communicative and strategic action thus betrays a metaphysical impulse "to separate intrinsic from extrinsic or pure from corrupt and deem the latter irrelevant" or "unworthy of separate consideration".³¹ Like all such impulses, it is best subject to a deconstruction which shows up the ever-receding interplay of the centre and the margin, and the arbitrary privileging of the former.

To counter this objection, Habermas can reply that what counts as 'privileged' in his account of linguistic interaction is by no means arbitrary. For if language is to play its role as the mediator of the (even creative) subject, it must be able to co-ordinate the everyday interactions of subjects. And the co-ordination of actions which is a necessary condition of social life - and therefore of an individual's creative linguistic life - is itself conditioned by the presupposition of intersubjectively identical ascriptions of meaning in the content of speech-acts. This shared consensus defines the literal meaning of the speech-act; the meaning which is exploited by an illocutionary force and which emerges under the constraints of rule-bounded action-coordination.

It is thus the *action-coordinative* property of ordinary language which is at the root of the asymmetry between it and its poetic/fictive derivatives. These derivatives gain their power partly from the bracketing or withdrawal of the illocutionary force with which they are normally deployed for the sake of action co-ordination in everyday life. This force can be bracketed where speech-acts are relieved of the pressure of action-coordination. Even in contexts where the particular illocutionary force of an utterance is bracketed, as for instance with a quoted promise in a play, its meaning still depends on the existence of conventions which co-ordinate the actions of the actors qua actors - as speakers and hearers with a determinate kind of function which is mutually recognized. The effectiveness of the quoted or reported promise therefore presupposes "the constraints under which illocutionary acts develop a force for co-ordinating action and have consequences relevant to action" in the realm of ordinary or 'normal' language.³² It is just these "consequences relevant for action" which specify what Habermas means by "standard" conditions.

Habermas is surely right to assert that the imaginative, creative linguistic inventions of irony, metaphor, and so forth, could not alone co-ordinate actions in the manner required for the social reproduction of life. In other words, they are incapable of performing that socially integrating operation which is the focus of Habermas's theory to the extent that it is *acultural* in kind. They *are* 'irrelevant' from the acultural perspective of Habermas's theory, but Habermas does not dismiss them as unworthy of *separate* consideration. Habermas repeatedly insists on the importance of tapping the critical potential of the imaginative, creative, linguistic activities informing modernist art. But rather than having the *acultural* significance accorded to 'problem-solving', action-coordinative mechanisms, this kind of linguistic activity serves what Habermas calls a "world-disclosive" function. Though introduced by Habermas only after the composition of *The Theory of Communicative Action*, this distinction between the problem-solving and world-disclosive functions of language is crucial for understanding the role of communicative action in Habermas's proposal for the self-reassurance of the project of modernity.³³ Communicative actions gain their c-ordinative property by virtue of the bonding established by the warranty to justify the validity claim raised by the illocutionary speech act. To this extent, they meet problem-solving requirements which any society must satisfy. The world-disclosive function, on the other hand, releases the speech-act from illocutionary obligations and problem-solving contexts of everyday life. For Habermas, the deconstructivist focusses narrowly on this latter function, but it is only possible in a language which already proves its worth in the problem-solving contexts of everyday interaction.

But what does it mean to say that a language must 'prove its worth'? The significance of

this point can be drawn out of an objection Habermas makes of deconstructive practice in general. Any interpretative practice, Habermas rightly asserts, must make intelligible its own possibility of communication. An interpretative practice premised on the idea that "every reading is also a misreading" fails this test, Habermas claims, in denying criteria for judging between interpretations and misinterpretations. While we can agree with this counter-objection to deconstruction, Habermas *then* goes on to claim that no matter how far removed interpretations are from the restraints of the everyday communicative situation, "they can never be wholly absolved of the idea that wrong interpretations must in principle be criticizable in terms of a consensus to be aimed for ideally."³⁴ The 'proving of worth' of ordinary language, by implication, resides not in its *de facto* meaningfulness; it is not by appeal to linguistic conventions that standard conditions are established. "Rather", Habermas continues:

language games only work because they presuppose idealizations that transcend any particular language game; as a necessary condition of possible understanding, these idealizations give rise to the perspective of an agreement that is open to criticism on the basis of validity claims.³⁵

Language games are therefore in continual need of justification - having 'to prove their worth' and 'subject to ongoing test' - in terms of such an ideal consensus. Hence, the satisfaction of these demands in everyday practice is what justifies the derivation of 'parasitic' from 'normal' language use.

The deconstructive criticism about the vagueness of normal or standard conditions has not been fully met, since the deconstructivist might rejoin that the selection of *which worth* is proved only pushes back Habermas's theoretical bias. But in the terms of the debate which

we are exploring, it is not - as the deconstructivist would claim - the theoretical bias *qua theory* which is challenging in this criticism, but its bias *qua acultural*.³⁶ The dispute between Habermas and deconstruction on the role of illocutionary force thus brings us to a central tenet of Habermas's acultural critical defence of modernity. The rationally motivated *ideal* consensus offers a 'culture-neutral' critical perspective on modernity, while being conditioned by the differentiation of validity claims which characterizes modern, problem-solving learning processes. It is therefore qualified as the basis of a critical self-reassurance of modernity out of itself. But the idealizations required for valid problem-solving do not enter into the 'world-disclosive' capacity of language, since it is not amenable to the kind of ideal consensus which Habermas takes to be presupposed in the ascription of identical meanings required for sustainable linguistic action-coordination. It follows that Habermas is unable to accommodate the evolution of world-disclosive powers within his theory of communicative rationalization, even though - as I will argue in the next chapter - he implicitly relies on it in his Hegel-derived plea for critical self-reassurance.

(2) *The Agonistic Objection*: The agonistic objection shares with the deconstructivist one a suspicion that the imaginative use of language is occluded or unduly marginalized in Habermas's emphasis on the consensus-building (or consensus-presupposing) function of language. But it proceeds not so much by stressing the interdependence of communicative and strategic linguistic interaction, as by prioritizing the strategic dimension of language as 'the first principle' of understanding the linguistic social bond.³⁷ Accordingly, for Lyotard, "speech acts fall within the domain of a general agonistics", which is to say within the domain of strategic interactions of adversaries.³⁸ This is of aesthetic significance not only because one of the chief adversaries is "the accepted language", but also because it is

strategically effective to create new 'moves' rather than relying on "reactional countermoves", which Lyotard suggests are "no more than programmed effects in the opponent's strategy".³⁹ According to Lyotard, speech acts are best accounted for as moves made between opponent players of language games in the medium of which social bonds are forged.

The social bond is considered by Lyotard to be, at the very least, a function of language 'effects'. These effects position the sender, addressee, and referent of speech-acts, as nodal points within the perpetually shifting local networks of strategic relationships which constitute the social. The ammunition for the war which is ordinary language is provided by these effects, the diversity of which encourages the highest degree of flexibility of utterance in everyday discourse. But this flexibility becomes ossified by institutional constraints which privilege particular kinds of language game; "orders in the army, prayer in the church, denotation in the schools, narration in families, questions in philosophy, performativity in business."⁴⁰ Yet this tendency towards the "bureaucratization" of the social bond, Lyotard suggests, is simultaneously threatened by the possibility that such institutionally imposed limits on potential language moves be themselves taken as "the stakes and provisional results of language strategies".⁴¹

For Lyotard, then, the way to approach speech acts is as moves in language games made between adversaries according to strategies with the power to subvert established meanings. This approach, he thinks, shows greater sensitivity to the diversity of language games which people play - games which have no metalanguage to commensurate them. But to accept this incommensurability is to reject the consensus model of linguistic interaction, since it is to

deny the assumption that;

it is possible for all speakers to come to agreement on which rules or metaprescriptions are universally valid for language games, when it is clear that language games are heteromorphous, subject to heterogeneous sets of rules.⁴²

Consensus cannot be the goal of dialogue, Lyotard insists, since there is no possibility of a metalanguage into which the diversity of language games could be translated and in terms of which agreement could be formulated. And worse, consensus reinforces that tendency towards bureaucratization - towards the ossification of language and hence of the social bond - by imposing the very conformity which is resisted by the strategic tapping of the heterogeneous potential of language games.

One can say then that from Lyotard's agonistic perspective, the distinction between "manipulatory speech" and "free expression and dialogue" is to be rejected on three counts.⁴³ First, because all utterances have effects which are of significance in the forging of social bonds. Second, the distinction overlooks the diversity of effects specific to utterances within different language games, of which Lyotard mentions "denotatives", "prescriptives", "evaluatives", and "performatives". And third, because "free expression" can itself be considered as creative manipulatory intervention oriented towards the dissolution of consensus.

Habermas might reply to these objections as follows. Regarding the first point, although it may be true that utterances are constantly placing and displacing sender, addressee, and referent in such a way as to make and break social bonds, we can nevertheless distinguish

between those bonds which are forged on the basis of the *validity* of the utterance, and those which are otherwise motivated. Habermas may exaggerate the *degree* to which the bonding between individuals in modern societies is ascribable to this kind of rationally motivated agreement, but the agonistic objection fails to account for the very possibility of it. This relates to the second point, since the co-ordinating effect of the rationally motivated agreement depends upon the recognition of the distinct *validity-claims* raised in the utterance. So the "effect" of a denotative will differ from that of a prescriptive for no other reason than that for which the effects of a communicatively and strategically oriented denotative utterance differ, except insofar as different resources of the lifeworld are at stake in the communication. As Peter Dews has observed, since these resources correspond to the different kinds of validity-claim, Lyotard's objection appears from the Habermasian perspective as a confusion between language-games and validity-claims.⁴⁴ Denotatives do indeed differ from prescriptions, but *qua* denotatives and prescriptions differ not in their "effects" (which depends on the context of the speech-act), but on the validity-claim (in these cases truth and normative legitimacy) that they *thematize*.

It follows that the charge against the imputed assumption of a metalanguage by virtue of which all different language games are commensurable is misplaced. Validity claims are already either implicit or explicit in the diversity of language use, in such a way that no further commensurating metalanguage need be invoked to account for consensus about them. What is perhaps more worrying about Habermas's position, however, is that it does seem to require that all claims to validity are differentiable into truth, normative legitimacy or rightness, and sincerity. The worry is that there may be validity-claims the truth-evaluability of which is inseparable from their action - orienting legitimacy, and not just

because these two aspects of validity are raised simultaneously. In the next chapter I will argue that this holds for a key class of critical concepts; those which feature in what Habermas calls 'clinical intuitions'.⁴⁵

Turning to the third aspect of the agonistic objection, is language oriented to consensus disposed to 'bureaucratization' in the manner Lyotard suggests? The de facto agreements concerning obedience to military orders, rote learning in school, and efficiency in business have no bearing on the issue, of course, because the kind of consensus proposed by Habermas is an ideal one, the conditions of which are not met where institutional constraints are normatively ascribed rather than communicatively established. Habermas can reply here that the objection rests upon an equivocation concerning the concept of 'strategy'. To be sure, the institutional ossification alluded to by Lyotard can be strategically subverted by invention, but this is just to say that a condition of communicative action is the setting loose of all three validity claims. Indeed, the scenario depicted by Lyotard represents on the one hand systemic constraints upon communication, and on the other hand, a one-sided rationalization of lifeworld institutions. The tendency towards bureaucratization, therefore, is just as well explained internally to Habermas's position. And it could further be argued that the strategy of subversion is better explained within it, since it gives a *point* to the subversive critique beyond the sheer strategy of subversiveness - namely, in the establishment of communicative interaction *in its full scope*.⁴⁶

Habermas agrees with Lyotard that artistic creativity has the critical potential to let loose identities and social structures ossified by bureaucratic institutional constraints. I will return in chapters five and eight to the problems Habermas has in grounding this claim, but now

I want to move onto an alternative argument to support his thesis that strategic linguistic interaction is in principle derivative from linguistic interaction oriented towards consensus.

(4.4) The Primacy of Communication: An Alternative Argument

I indicated when considering the deconstructivist objection to the primacy of 'reaching understanding' Habermas's view that communication is possible "only under the presupposition of intersubjectively identical ascriptions of meaning". The possibility of perlocutionary or teleological success presupposes the prior possibility of reaching understanding in that the utterance employed strategically, to be effective at all, must first be intelligible to the opponent. In order for the manipulation of meanings to be possible, there must first be meanings to manipulate - there must be something to be used as a means to an end. But if meaning or intelligibility is itself not something which can be strategically decided upon, then the strategic use of language cannot be originary.

Habermas recognizes this point, and interprets Wittgenstein's remarks on rule-following, which are pertinent for Habermas's purposes in explaining what makes for the 'sameness' of the various applications of a meaningful concept, in a way which indirectly supports his own thesis that the communicative use of language is the original mode.⁴⁷ But it is just at this point that the Wittgensteinian counter-objection can be wedged. For although Wittgenstein gives reasons for thinking that some form of consensus is a necessary condition of meaningful utterance, it is questionable if the consensus which must be presupposed in acts of communication is of the *kind* proposed by Habermas, and at the *level* he takes it to be operative. And it is just the kind of consensus and the level at which communication presupposes it which are the crucial features of the moral Habermas draws for his model

of pure communicative action, intact intersubjectivity, and critique. In the rest of this section, I shall rehearse the argument Habermas draws from Wittgenstein's remarks on rule-following, and indicate how it serves to bolster the thesis that the communicative use of language is originary. In the next section, I shall consider how the same remarks also serve to undermine that thesis.

For Habermas's purposes, the import of Wittgenstein's remarks on rule-following rests in establishing an analytical connection between "identical meanings and intersubjective validity"⁴⁸. Rule-following is the *sine qua non* of propositionally differentiated language use since it determines what counts as the correct application of a concept. Of course, if there were no such thing as the correct application of a concept, there would also be no such thing as an incorrect application. And if this were the case there would be no concept to be applied or misapplied at all. This is the objection Habermas puts to the deconstructivist. If the meaning of a concept is to be the same in the ongoing applications of it, it must be so by virtue of a rule which determines what is to count as the 'same' meaning. Habermas rightly points out that this 'sameness of meaning' is not something which can be inductively inferred from empirical regularities in the application of the rule, since the rule itself is needed to determine what *counts as* a particular instance of an empirical regularity. For this reason, what counts as 'going on in the same way' in the application of a rule cannot be determined by appeal to something *external* to the rule and its application.

The ability to follow a rule is the ability to apply the rule in the same way to different particular cases. But what counts as the same way is not reducible to any empirical or *non-normative* phenomenon. Habermas rightly takes this to imply that a condition of

'sameness of meaning' is intersubjective validity in the application of concepts. Given that identity of meaning is conceptually tied to intersubjective validity, it follows, according to Habermas, that the violation of a rule by a particular subject must be criticizable by another subject who has grasped the rule. Further, this will be by way of a "critique which is in principle open to consensus".⁴⁹ "Without this possibility of reciprocal criticism and mutual instruction leading to agreement", Habermas continues, "the identity of rules cannot be secured. A rule has to possess validity for at least two subjects if one subject is to be able to follow the rule".⁵⁰ This is a consequence of Wittgenstein's argument against the possibility of a subject following a rule 'privately'. The intersubjective validity of rules is a validity which must obtain between at least two subjects where both "must have a competence for rule governed behaviour as well as critically judging such behaviour".⁵¹ A single isolated subject, who could either not act in accord with the rule or critically judge rule violations, "could no more form the concept of a rule than he could use symbols with identically the same meaning".⁵²

For Habermas, then, sameness of meaning, being bound to the intersubjective validity of rules, implies the ability to take a "critical yes/no position" on the correctness of a particular application of a rule. But such a concept of rule-competence refers, inter alia, "to the ability to produce symbolic expressions with communicative intent and to understand them".⁵³ So at the most fundamental level, linguistic competence presupposes the grasping of rules the validity of which is in principle open to a consensus on the basis of "reciprocal critique" and "mutual instruction". The qualifiers "reciprocal" and "mutual" are crucial here, since they circumscribe the critique and instruction presupposed in the very foundation of language use within the exclusive realm of the *illocutionary*. And implicitly, the kind

of consensus to which the intersubjective validity of rules is accountable is one divorced from causally induced effects. Since any actual consensus is always conditioned by some empirical motivation, there is an *ideal* consensus presupposed in the original use of language. Although an idealization, it is nevertheless an assumption which the possibility of identical ascriptions of meaning demands. Virtual to linguistic competence, then, is the consensus of an ideal communication community, which is to say the consensus of subjects in a dialogue situation pursuing solely illocutionary aims - or in other words - pure communicative action.

Of course, by this point Habermas recognizes his radical deviation from Wittgenstein's position. To return to Habermas's reply to the deconstructivist objection, and now also *contra* Wittgenstein, Habermas insists that it "is not habitual practice that determines just what meaning is attributed to a text or an utterance", but rather that "language games only work because they presuppose idealizations that transcend any particular language game".

Habermas's reading of Wittgenstein gives a distinctive and novel twist to the recent and voluminous literature within analytical philosophy on Wittgenstein's rule-following considerations. Outside this interpretative tradition, however, Habermas follows Apel in conceiving communicative discourse as presupposing an entwinement between "a real communication community" into which the participant has become socialized, and "an ideal communication community" in which arguments would be properly understood and judged.⁵⁴

As the previously quoted remark suggests, the idea of an ideal communication community is invoked to save both rationality and critique from the conservatism of habit and custom.

I will now consider whether this move is compatible with Wittgenstein's remarks on

rule-following from which Habermas draws:

(4.5) Consensus and Strategy in Language; A Third Set of Objections

Can Wittgenstein be "turned on his feet"⁵⁵ so as to show the primacy of communicative action and the immanence of the ideal communication community? Is linguistic interaction oriented to reaching an intersubjectively valid consensus shown to be presupposed by (or analytically connected to) the possibility of identical ascriptions of meaning? As I indicated, the problem lies in the kind of agreement which is presupposed in communication, and the level at which this agreement is operative. On the first point, Wittgenstein is emphatic that the kind of agreement which valid argumentation presupposes is not one of opinions, but of *practices* or 'forms of life'. The consensus here is not one of the 'yes' or 'no' of the interpreter of the rule. It is not an *interpretation* which is in principle open to a consensus, but the background against which interpretations are made. Following a rule is 'founded on agreement', but agreement over the background framework of action "within which the concept of following a rule has intelligible employment, not to the explanation of what 'following a rule' means".⁵⁶ So not only is the kind of agreement at issue here one of practices rather than opinions, but the level at which it operates is at the level of the *pre-conditions* of correct or incorrect opinions. To paraphrase a remark of Wittgenstein's to make it directly applicable to Habermas's view, "the agreement of validations is the pre-condition of our language game, it is not affirmed in it".⁵⁷

Habermas's fear is that this seems to abolish argument.⁵⁸ Argument - the raising, criticizing, and redeeming of validity claims - seems to be reduced to matter of fact habitual practices. But for Wittgenstein, neither actual nor ideal agreement is that by virtue of

which an application is in accord with a rule. No agreement is necessary to mediate between a rule and the correct application of it, since rule and application are, as Habermas recognizes, internally related. Indeed, Habermas recognizes and *affirms* many of the above points in his doctrine of the lifeworld. The following passage, in which Wittgenstein sums up the role of agreement in his rule-following considerations, highlights this affinity nicely;

It is of the greatest importance that a dispute hardly ever arises between people about whether the colour of this object is the same as the colour of that, the length of this rod the same as that, etc. This peaceful agreement is the characteristic surrounding the use of the word 'same'.

And one must say something analogous about proceeding according to a rule. No dispute breaks out over the question whether a proceeding was according to a rule or not. It doesn't come to blows, for example.

This belongs to the framework, out of which our language works (for example, gives a description).⁵⁹

What Wittgenstein refers to here as the 'framework' has just the same role which the 'lifeworld' has, *inter alia*, for Habermas. The question now is whether the primacy of communicative action is compatible with such a lifeworld concept. I will approach the nature of this tension by considering the Wittgensteinian set objections brought against Habermas's picture of critical reflection in a recent article by James Tully.⁶⁰

Tully's basic objection is that 'reaching understanding' cannot ground the certainties of everyday life, and that it is unreasonable to take the communicative action orientation to them. If Habermas were to utter the speech act "I am Jürgen Habermas and I believe that the workplace ought to be organized democratically", it would be unreasonable to ask for reasons concerning the sincerity of his claim about his name, and reasonable to take it for

granted.⁶¹ The example is trivial, but it is meant to show that rational action and belief often *involves* taking things for granted, rather than giving reasons which will ground their claim to validity; giving reasons and interpretations for following a rule comes to an end at the point where all one can say is that "This is simply what I do". But this would only undermine the rationality of the rule-following if some *foundational* relation were needed between the rule and the application, but the *internal relation* between rule and application offsets this requirement. The absence of a reason for applying a rule does not reveal an irrationality in the nature of rule-following; "to use a word without justification does not mean to use it without right", unless a justification or explanation is required by *us* "to prevent a misunderstanding".⁶²

But it is only on the condition that this will not always be the case that giving reasons, justifications, and validations can get off the ground at all. Giving reasons requires an acknowledgement of a standard for reason giving, just as the practice of measuring requires a background agreement upon the standards of measurement. To return to Tully's example, the sincerity of Habermas's declaration of his name could be questioned, and validated by the presentation of his birth certificate. This is the standard against which such claims are validated. In exceptional conditions, there may be reason to doubt such standards, but there will come a point at which the questionableness of them *threatens the possibility of communication*.

It follows that the use of language oriented towards a "communicatively reached agreement" which "must be based *in the end* on reasons"⁶³ cannot be the original use of language. For the rationally motivated agreement of communicative action presupposes a prior agreement

about the standards of reason giving. A reason cannot be given for the justification of these standards, since they determine what is to count as the justified and unjustified use of words at all. To put it in another way, Habermas's view that linguistic interaction oriented to a validity conditioned agreement falls foul of Wittgenstein's point that not all propositions can be *problem-solving hypotheses*, since the very possibility of judgement presupposes commonly agreed standards or norms of judgement. If communicative action were originary, then it would be possible that all propositionally differentiated speech acts were hypothetical in form.

On the face of it, Habermas's reply would seem straightforward; these Wittgensteinian objections neglect that complementary to the concept of communicative action is the concept of the lifeworld. For the lifeworld, according to Habermas, is just that tacitly accepted and mutually agreed upon background framework of taken-for-granted assumptions and meanings against the horizon of which validity claims are criticizable. The problem remains, however, that the agreement to which 'reaching understanding' is oriented must be of a *different order* to the background agreement in action which constitutes the lifeworld. This suggests that the concept of the lifeworld serves more to undermine rather than to complement the concept of communicative action, at least insofar as the latter is understood as the original mode of language interaction.

Consider again Wittgenstein's remark that "agreement of ratifications is the pre-condition of our language game, it is not affirmed in it". Now the agreement to which communicative action is oriented, unlike the lifeworld agreement in action, is precisely such an affirmation. Habermas brings together these radically different senses of consensus in

his concept of communicative action, enabling him to shift from the original validity *conditioning* lifeworld agreement to the validity *conditioned* communicative agreement, and then to assert the latter as the originary form of agreement or consensus which all linguistic interaction presupposes. Further, the level at which the fundamental agreement operates is not one which transcends the internal relationship between the rule and its correct application. Such an idealization is thus not necessary to *explain* intersubjectively identical ascriptions of meaning, but it is just upon this presupposition that the primacy of communicative action thesis rests.

But although we can now say that communicative action is not the originary form of linguistic interaction, we are also committed to saying that *strategic* interaction is not fundamental *either*. Background lifeworld agreement is not a 'giving something to understand' nor an inconspicuous harnessing or manipulation of another person's intentions for one's own purposes. Both strategic and communicative action as defined by Habermas imply a voluntarism which the coherence of lifeworld agreement will not tolerate. We can see a way out of this incoherence only by returning to the question initially posed by Habermas and taken up by his critics: it asks for an original mode of language use, to which communicative and strategic action are proposed as answers. That neither is adequate suggests that there may be something misleading in the formulation of the question itself. Is the idea of an 'original use' of language really intelligible at all? Within Habermas's argument, 'use' stands in for *intentional* action, and he considers two different kinds of action-orientation as possibly originary. But the concepts about which action is oriented cannot themselves be decided upon intentionally, *whatever* the orientation of the action. It therefore makes no sense to talk of an original mode of language use, if by

language 'use' is meant a kind of intentional action. Both communicative and strategic action presuppose the possession of concepts which are prior to any intentional use of language.⁶⁴

(4.6) Conclusion

While communicative action may be conceptually prior to strategic action to the extent that language is a mechanism for the co-ordination of intentional actions, Habermas has not shown that this is something which can be established on the basis of the intelligibility of the domain of meaning as such. For Habermas, this domain is opened up by the *pragmatic* presuppositions of communication, by the illocutionary bonding force of the warrant to satisfy what I called in chapter two the 'validational imperative' over the whole range of propositionally differentiated utterances. By arguing for the originary status of communicative action, Habermas has sought to make space for a claim to normativity in the *form* which the pragmatics of communicative competence unavoidably imposes on language users. But the space for rational normative critique which is promised by the pragmatic presuppositions of language in its originary action-coordinative mode can seem to be threatened by the Wittgensteinian objections. For the Wittgensteinian position seems to "leave everything where it is"; it appears to give up on the question of the *accountability* of habitual practices. If linguistic practices are self-justifying, if there is no possibility of a standpoint which transcends them and from which they can be criticized, are we not left with a linguistic positivity of the present? And if this is the case, we seem to be no better-off in our pursuit of the self-clarification of the claim to normativity. But as yet, all that the Wittgensteinian position entails is that the self-clarification of critical reflection cannot decide in advance how meaningful criticism is "to go on" in its various contexts of

application. In the next chapter I will consider how Habermas's construal of what it is "to go on" in the practice of critical reflection concerning the modern project clashes with the model for the need for self-reassurance he derives from Hegel's idea of a 'dirempted' ethical totality.

CHAPTER 5: CRISIS AND CRITIQUE: TENSIONS IN HABERMAS'S CONCEPTION

(5.0) Introduction

I have been taking as fundamental Habermas's contention that a standard of 'intact intersubjectivity' offers itself to a critical philosophy which reflects upon the communicative mediation of human beings. It remains to clarify the measure of this standard, and to assess the scope of its critical powers. I will begin (5.1) by considering how Habermas distinguishes between the basic types of phenomena which motivate critical reflection upon modernity. These peculiarly modern phenomena, Habermas maintains, are criticizable only according to the standard which is appropriate to their type. Unfortunately, however, Habermas uses the *same* terms - 'intact intersubjectivity' and 'ethical totality' - to refer to these different standards when dealing with the problem of self-reassurance. At this point, I identify three different - though closely related - senses of Habermas's concept of an 'ethical totality', which allows me to argue subsequently that Habermas trades on this ambiguity in order to resolve tensions between the Hegelian concept of crisis considered in chapter three, and the validational concept of critique discussed in chapter four.

In (5.2), I indicate why, in accordance with the different standards available for critical reflection, Habermas stresses that the undertaking of the critical self-reassurance of modernity must obey a strict division of labour in its method. On the one hand, there is what Habermas calls the "rational reconstruction" of stages of competence; on the other, there is what he variously calls "methodically carried out self-critique" ¹, "interpretation on behalf of the lifeworld" ², and "mediation on the part of the lifeworld" ³. I will then go on

to question the viability of this division of critical labour by arguing, in the core section of this chapter (5.3), that it leads to a distorted picture of practical reasoning. By taking the moral domain as suitable for rational reconstruction, hence criticizable in virtue of a *procedural* standard of intact intersubjectivity, Habermas understates the scope to be covered by the self-clarification of the claim to normativity. This charge need not worry Habermas so long as he can show that adequate 'compensation' can be given to the task of the rational reconstruction of a narrowly circumscribed moral point of view. Habermas suggests that such compensation can be provided by the 'clinical' intuitions of what he calls 'therapeutic' reason. After arguing that 'compensation' is an inadequate term for grasping the relation between moral judgement and clinical intuition, I suggest that we need to reconsider the relationship between what Habermas calls practical and therapeutic rationality. This I do in section (5.4), where I trace an insuperable difficulty facing Habermas's conceptualization of the rationality of clinical intuitions - of the criteria of sickness and health which would provide a *substantive* standard of intact intersubjectivity - back to acultural theoretical commitments which inform the rational reconstruction model of critical reflection. I propose that this difficulty could be overcome if Habermas were to abandon some of these commitments, and in the conclusion I consider two reasons Habermas puts for not rejecting them. These reasons are seen to require further substantiation.

(5.1) Diremption, Differentiation, and Disharmony

There are three analytically distinct, though causally related modern phenomena, which contour the object-domain of critical reflection for Habermas. Most important, there is the 'diremption' which Habermas calls the 'uncoupling' and subsequent 'colonization' of the

lifeworld. The key here, as I explained in (1.2), can be put as follows. Individuals are always mediated by the process of socialization. The stability of societies, and therefore also of individuals, turns upon the degree to which the identities of individuals (and collectives) can be secured in the process of social integration. In pre-modern societies, religion provided this integrating force or 'unifying power'.⁴ Identities could be secured in an intersubjectively binding way through the conviction of a cosmically-realized moral order. With the unfolding of modernity, Habermas's thesis runs, the unifying power of religion is weakened irredeemably. Increasingly, modern societies come to rely upon communicative action for the reproduction of their symbolic resources: as societies evolve, the unifying or bonding force which integrates them feeds increasingly off the force of the better argument. By the 'force of the better argument' is meant the speech act obligations immanent to a rationally motivated agreement, rather than empirically determined, de facto accord. I have already considered some of the general conceptual problems attending this idea of a rationally motivated agreement, but for the moment we are concerned with the costs of the modernizing process, assuming for the sake of argument that Habermas's description of it is accurate.

First, Habermas recognizes that as a vehicle for forging social bonds, the speech-act immanent obligations of communicative action are extremely precarious. The modern subject, as mediated by the risky self-formative process of communicative action, is constitutively fragile and vulnerable.⁵ But this fragility can turn to collective *pathology* when the communicative mediation of subjects is systematically distorted. For alongside the rationalization of its socially integrating process - and thus of its mechanism for reproducing symbolic resources - modernity develops its powers of material production

according to the logic of capitalist systemic growth. Under the material imperative to maintain itself as a system steered by money and power, modernity's already fragile social integration faces a more serious threat. The force of the better argument can give way to bureaucratic and market forces which cannot sustain the identities of individuals and collectives. A communicatively mediated lifeworld colonized by a strategically mediated system then avenges itself in the crises and social pathologies distinctive of modernity. Diremption refers fundamentally to the crises generated by the splitting-off of areas of the communicatively integrated lifeworld by the functionally integrated economic and bureaucratic system. Crises are generated because such a *norm-free* mediation of subjects cannot secure the identities of individuals nor bind them together in solidarity as collectives. Habermas sometimes uses the concepts of the 'lifeworld' and 'ethical totality' co-terminously,⁶ and I shall call the ethical totality *qua* lifeworld (the source of normativity colonized by the norm-free steering media of the system) 'ethical totality'.

The critique of diremption, then, takes as its object the costs of societal rationalization, reconstructed according to the kind of *logic of development* outlined above. This phenomenon, Habermas claims, corresponds to what Hegel described as the 'positive'. Its critique goes by way of reconstructing the unifying potential of communicative rationality. But while, considered as a resource, the lifeworld is concrete and historical, the unifying force of communicative reason is formal and procedural to the extent that the intersubjective bond forged through it is based on a warranty that norm-validation can be provided on demand independently of the content of identity-carrying convictions and beliefs. As we saw in the previous chapter, Habermas distinguishes the unifying, socially integrative force of communicative rationality from the 'causally induced' bond established by 'empirically

motivated' assent. But since all *actual* agreements are empirically motivated, Habermas refers to this standard for critical reflection or source of normativity as the counterfactual or *ideal* speech situation. As participants in communicative action must presuppose such a situation obtains *qua* rationally motivated, it can also be considered as *virtual* to all cases of communicative action: "sociocultural forms of life stand under the structural restrictions of a communicative reason *at once claimed and denied*".⁷ These *structural* restrictions define the formal standard of intact intersubjectivity. Let us call this formal, counterfactual standard of an ideal speech situation or 'ideal communication community' 'ethical totality'.

The critique of *cultural* rationalization, as I introduced it in (1.1), has a quite different object-domain. Culture, for Habermas, becomes rationalized as reason gets differentiated into its theoretical, practical, and aesthetic dimensions. Habermas does not take this differentiation - reflected in the separation of the cultural value spheres of science and technology, law and morality, and art, and the relatively autonomous development of discourses thematizing validity claims of truth, justice and taste - as in itself a source of discontent. But it does make a modern source of discontent possible, since 'expert cultures' are able to separate themselves from the mainstream of everyday life. Questions of scientific truth, legality, and aesthetic worth become increasingly subject to ever more complex professional assessment. This generates a two-fold problematic. First, within a particular rationalized value-sphere there arises the problem of how those aspects of validity not thematized within it are to exercise their force. Second, there is the problem of how the knowledge which is creamed off by the expert cultures across all the value spheres is to be allowed to feed back into a culturally impoverished everyday practice. Both are problems of the *mediation* of what is necessarily differentiated in the rationalizing process.

Habermas takes them to be problems which are independent of the process of societal rationalization; they have no bearing on the *structural* pathologies generated by diremption.

It follows that the standard for critique in this domain cannot be the formal standard of an intact intersubjectivity. For even where the structures of intersubjectivity are not damaged, the same problems of cultural impoverishment can arise. Habermas makes only a passing reference to what standard of critique can be brought to bear in dealing with these problems of mediation. He speaks of a "free interplay" between the separated moments of reason which have come to a standstill like a "tangled mobile".⁸ I will comment briefly on his suggestion in section (5.2), and in (5.3) I will propose that the 'problem of mediation' both within the differentiated and rationalized value sphere of morality, and between expert moral insight and everyday life, runs deeper than Habermas's understanding of differentiation leads us to suppose.

Habermas also employs this standard of a 'free interplay' to address a problem of mediation which *is* connected to diremption. For under the pressure of systemic growth, the modern lifeworld becomes selectively exploited for its cognitive/instrumental rationality potential. That is, it suffers a *one-sided* rationalization which is ultimately to be explained in terms its splitting-off or diremption by the system. Hence the overall object of critique is formulated as the "division [diremption] and usurpation [one-sided rationalization]"⁹ of a communicatively mediated, procedurally unified lifeworld. So for his standard of intact intersubjectivity against which deformations of forms of life *as a whole* is to be measured, Habermas appeals to a notion of the free interplay of the three different dimensions of communicative rationality: an everyday practice which is "open to an *uninhibited* and

balanced interpenetration of cognitive interpretations, moral expectations, expressions and valuations".¹⁰ Let us call this standard of intact intersubjectivity 'ethical totality'.

In reply to the objection that *communicative rationality* does not avail itself of such a standard - an objection I will return to from (5.3) to the end of this chapter - Habermas suggests that the pathological symptoms of imbalance can be explained in terms of systematically distorted communication, formulated in terms of an "inflexibility" when the internal links between meaning and validity, meaning and intention, and meaning and accomplished action are interrupted.¹¹ By implication, such a standard of an intact intersubjectivity would not be a merely formal notion, since it has *substantial* implications for the assessment of the well being of forms of life as a whole. However, Habermas prefers to leave the criticism of the *disharmony* of individual and collective life forms, the space it leaves for the passage of an "undamaged, correctly spent life", to the "clinical intuitions" of the critic. These intuitions give content to a standard of intact intersubjectivity against which the integrity of forms of life as a whole can be judged. Habermas insists that this is not to be confused with the formal (or procedural) standard of intact intersubjectivity qua the ideal speech situation. The latter, we are warned, does not serve as the "image of a concrete form of life".¹² But in what sense, if any, can such clinical intuitions be rationally justified, if rationality is defined procedurally as communicative competence?

If Habermas cannot successfully address this question, the consequence for his contribution to the philosophical discourse of modernity is potentially devastating. As he formulates it, this discourse addresses the problem of modernity's critical self-reassurance - of how

modernity can not only generate its norms out of itself, but also *stabilize* itself on this basis. But Habermas is also committed to the acultural theoretical requirement of explaining the transition to modernity as process of rationalization. So the question he has to answer is how normativity is generated in the process of the rational mediation of subjects. He does this by appeal to the presuppositions of the distinctively modern vehicle for rationally reproducing symbolic resources (communicative action), which he captures in the formal concept of an ideal speech situation or 'ethical totality'. But as we saw in (3.1), he also explicates Hegel's idea of a dirempted ethical totality which inflicts "suffering due to alienation" as a "violation of the structures of *rational* life together".¹³ He must therefore account for how the disruption of 'ethical totality' can be theorized as *irrational*. In other words, he must show how 'clinical intuitions' concerning the well being of modernity as a whole can be derived from the norm-carrying medium of communicative action - if modernity is to have grounds for self-reassurance. Alternatively, the normative basis of these clinical intuitions could be re-theorized *outwith the constraints of an acultural theory* which would derive their rational core from the culture-neutral mechanisms of symbolic reproduction and social integration. My argument will be that only the latter move can save Habermas's model for self-reassurance from the incoherence which this acultural commitment threatens to bring to it.

(5.2) The Divided Labour of Critical Reflection

In (1.1), I offered a sketch of the acultural theoretical strategy adopted by Habermas in order to explain the transition to modernity as a process of rationalization. We saw that the explanation appeals to "internally reconstructible sequences of stages of competence" applied to two cultural invariables; the reproduction of symbolic and material resources.

In (1.3), I introduced Habermas's conception of critical reflection in counterpoise to Gadamer's 'universality' claim for hermeneutics. The advantage which Habermas claims for his conception of reflection over Gadamer's, we saw, lies in its capacity understand, explain, and thereby overcome, phenomena of systematically distorted communication. This was to be the achievement of a theory of communicative competence. In (3.2), it was shown that this process of understanding, explanation, and overcoming is equivalent to an articulation and disenchantment of what Hegel called the 'causality of fate'. For Habermas, the causality of fate represents the dynamic of of a 'dirempted ethical totality' which is the hallmark of modern crisis-phenomena and the source of its need for self-reassurance. We saw how the dynamic of disturbance and reconstitution in the individual's psycho-sexual development - and therefore a self-formative process - relies on the moral responsibility assumed by the agent in her act of self-reflection. As Habermas put it in *Knowledge and Human Interests*, the resolution of the identity-crisis into which the agent is plunged through systematically distorted communication requires a 'passion for critique'.

But following upon the objections made to his early outline of the tasks of critical reflection in *Knowledge and Human Interests*, Habermas has been emphatic in distinguishing - *within* a theory of communicative competence - between the philosophical tasks of 'rational reconstruction' and 'methodically carried out self-critique'.¹⁴ These philosophical tasks, according to Habermas, best fit a post-metaphysical, non-foundationalist paradigm of critical reflection. Rational reconstructions attempt to articulate in a theoretically convincing manner the pre-theoretical know-how implicit in competent speech and action. This competence is assessed in terms of the intuitive mastery of rule systems. The command of these rule-like procedures is presupposed in the ability to produce such things as correct

inferences, good arguments, grammatically correct sentences, successful speech-acts, effective instrumental action, and appropriate moral judgements.¹⁵ Since what is reconstructed are formal procedures which condition claims to validity, Habermas construes this task as tantamount to the production of a general theory of rationality. The reconstructive task of Habermas's conception of critical reflection assumes a constructive role in providing strong universalistic but fallible knowledge claims, thus heightening the self-consciousness of subjects capable of speech and action. It assumes a *critical* role insofar as rational reconstructions explain deviant cases of incompetence. Although their domain is that covered by disciplines like epistemology, philosophy of language, and ethics, they are essentially strong empirical theories for which philosophy is a 'stand-in'.

Habermas's underlying thesis is that a standard for the critical self-reassurance of modernity can be generated out of distinctively modern forms of interaction. Given these forms of interaction, the argument seeks to establish that certain normative constraints - by appeal to which critique can proceed and which supply modernity with its normative content - must be presupposed by actors competent in them. As such, the argument is transcendental in type: it seeks to disclose the conditions under which a certain kind of practice is possible, by way of reconstructing the competences which are presupposed in it. But the claims of rational reconstructions are transcendental only in a weak sense; as part of a 'post-metaphysical' paradigm of reflection, they do not have *a priori* status. They are allegedly open to empirical refutation, and are always proposed with a fallibilist proviso; they do not have the epistemic certainty which could qualify them as 'foundational'.

The lack of a priori status of rational reconstructions is evident, as Benhabib has shown,

in their failure to establish the *necessity* and *uniqueness* of what they conclude as the implicitly known rules of rational speech and action.¹⁶ Any putative necessity is clearly undermined by the fallibility and revisability of their empirical support. But more crucially, the *uniqueness* of the rules which are claimed to be presupposed by a certain kind of action does not follow even from an argument which can establish such presuppositions. For there may be *other* presuppositions which are equally consistent with the 'given' that they condition. Further, these alternative presuppositions may be part of a *competing* framework of explanation. This may well mean that; on the one hand, at most one of the competing frameworks can be true - but also on the other hand, that if one is true that cannot itself be definitively established by this kind of transcendental argument.¹⁷ This thought suggests that the mode of verification of rational reconstructions may not be as dissimilar from the evaluation of frameworks of self-interpretation as Habermas supposes. This is a criticism I will explore further in the next section, where Habermas's application of rational reconstructions to competence in moral judgement is examined. But for the moment, what interests us is the contrast Habermas draws between the modes of critical reflection available to the theorist of communicative competence.

While rational reconstruction takes as its object the formal capacity for rational action and the logic (ie. bare or formal possibilities) of development of those capacities, methodically carried out self-critique employs narrative tools to make sense of the *dynamics* of development of the particular totality of an individual or collective way of life.¹⁸ By *methodically* carried out self-critique, Habermas implies that it should take the narrative but dialogical form for which the model of psychoanalysis is paradigmatic. Indeed, psychoanalysis is cited as a model for how rational reconstruction and self-critique can be

combined in the same critical project. However, only the former kind of reflection, for Habermas, can contribute to a theory of rationality. As we will see in the next section, this is because a claim to universality is ascribable to the competences reconstructed.

Habermas contrasts the role of reconstructive science with philosophy's role as interpreter on behalf of the lifeworld. As we have seen, in its latter role philosophy has no formal standard against which irrationalities can be gauged. Its task is to gather together or reconcile, at the level of everyday life, the moments of reason which communicative action tears asunder. As a reconstructive science, the theory of communicative action articulates the differentiation of the three validity claims and corresponding value spheres which Habermas accords to an evolutionary process of learning. But as interpreter of the lifeworld, it must show how mediation is possible between these differentiated dimensions of rationality. For Habermas, as we saw, the problem here lies not with the differentiation of the moments of reason as such by competent communicative action, but the way in which these moments become stuck like a 'tangled mobile'. In its latter role as interpreter or mediator on the part of the lifeworld, critical reflection must address the question; "How can a new balance between the separated moments of reason be established in communicative everyday life?".¹⁹ To the extent to which a new balance is called for by the one-sided rationalization of the lifeworld, which in turn is explained by diremption and subsequent colonization, the answer would provide a standard of 'intact intersubjectivity' which, in distinction from Habermas's other standards for critique, I earlier designated ethical totality^c.

(5.3) Morality and Ethical Life

The division of philosophical labour between rational reconstruction and interpretation or mediation on behalf of the lifeworld has deep implications for Habermas's understanding of ethics. The task of rational reconstruction, as we have seen, is to raise the intuitive know-how presupposed in the varieties of communicative competence to the level of theoretical self-awareness. Communicative competence presupposes the capacity to participate in discourse; to raise, and justify when challenged, particular validity claims of truth, rightness, and authenticity. For a claim to count as valid, participants in discourse must presuppose that a rationally motivated consensus concerning it is possible. Habermas refers to this counterfactual presupposition of discursive conditions, in which positions are taken up solely on the basis of the rational force of the better argument, as an ideal speech situation. Though counterfactual, Habermas insists that it is an unavoidable presupposition of rationally motivated agreement. Since it is unavoidable, it also conditions claims to universality. In making explicit the procedural rules which condition claims to validity, rational reconstructions render intelligible the universal basis of validity claims. The task of moral philosophy, accordingly, is to explicate the conditions which make rationally motivated agreement concerning validity claims to normative rightness possible.²⁰ To do this is simultaneously to explicate the universality of such claims and to construct a theory of practical reasoning.

Such is the burden of Habermas's 'discourse ethics'.²¹ It seeks to reconstruct the rational basis of the strong (because universal) but minimal (because formal) constitution of the 'moral point of view'. By the moral point of view is meant the impartial perspective of participants engaged in practical discourse. Practical discourse is a formal procedure of

argumentation, through which participants seek to arrive at a rationally motivated consensus over the legitimacy of a norm. The strategy of discourse ethics is to derive the universality of the moral point of view from the pragmatic presuppositions of moral argumentation. The thread of the Habermas's argument is as follows. The point of engaging in moral argumentation would be lost if it were not for the possibility of reaching a valid consensus. But the condition of reaching a valid (rather than de facto) consensus is that each participant has an equal right to raise criticizable validity claims, which is reciprocated amongst all the participants. From these procedural normative constraints on *participating* in practical discourse, a moral principle which is universally binding on communicative actors is derivable; namely, "For a norm to be valid, the consequences and side-effects of its general observance for the satisfaction of each person's particular interests must be acceptable to all".²²

What norms are valid can only be determined by participants in actual practical discourses. But to count as *just*, their content must pass the test of this formal principle of universalizability which defines the moral point of view. The moral point of view thus has a cognitive status; it is not contingent upon the particular cultural traditions or forms of life which give content to norms.²³ The moral sceptic is defeated, because a non-contingent, non-instrumental reason for acting is shown to have theoretical justification. According to the basic claim of Habermas's discourse ethics, the normative force of the pragmatic presuppositions of argumentation cannot be denied without a performative contradiction. Thus, the sceptical position is something which cannot be *argued* for, and if it cannot be argued for, it is not a 'position' at all.

This is what discourse ethics *qua* rational reconstruction attempts to demonstrate.²⁴ The object of reconstruction is the post-conventional moral consciousness of ideal-typical modern agency.²⁵ As Habermas acknowledges, however, limiting the domain of moral philosophy to the refutation of moral scepticism, and the reconstruction of the universality of the moral point of view, has its price. It can only be achieved if certain phenomena are excluded from the moral domain. Rather than pursue Habermas's formal-pragmatic refutation of moral scepticism, I want to focus on what is lost to ethical reflection once moral philosophy is understood as a stand-in for reconstructive science.²⁶ Even if the anti-sceptical conclusion could be established, it may be that its own point would be lost unless adequate *compensation* is given by the other critical role Habermas designates for philosophy - as interpreter and mediator on behalf of the lifeworld. But this is not what I shall be arguing. Rather, I will contend that the very issue of 'mediation' becomes *distorted* once it is construed according to Habermas's division of critical reflection.

Most conspicuously, Habermas's discourse ethics avowedly eschews consideration of what constitutes the good life. This is taken to be an inevitable consequence of a moral domain which is determined by a principle of universalizability. This principle "makes razor-sharp cuts between evaluative statements and strictly normative ones, between the good and the just".²⁷ Further, this separation between the good and the just is what conditions the possibility of cognitive advance through the exercise of practical rationality. It is just by virtue of the "transformation of questions of the good into problems of justice" that an autonomous moral domain owes its "gain in rationality".²⁸

One reason Habermas gives for this is that only questions of justice, or of the normative

validity of norms of action, can be debated with "the prospect of consensus".²⁹ Another is that evaluative questions, or questions of what constitutes the good life, "are accessible to rational discussion only *within* the unproblematic horizon of a concrete historical form of life or the conduct of an individual life".³⁰ Questions of the good life, Habermas continues, "have the advantage of being answerable within the horizon of lifeworld certainties".³¹ Those cultural values which make up one's conception of the good life can only be "candidates" for legitimate norms. They become objects of practical rationality, and potentially legitimate norms of action, as soon as a "hypothetical attitude" is taken towards them by participants in moral argumentation. Upon taking this attitude, the norms and institutions that are taken for granted appear as "instances of problematic justice".³² Particular norms can be tested as hypothetical legitimacy claims under the moralizing gaze of the problem-solving participants in practical discourse. And in its universality, this moral point of view stands *outside* the provincialism of the lifeworld.

The guiding intuition of Habermas's thought here is that the concrete institutions of any particular ethical context of life (*Sittlichkeit*) are criticizable by appeal to an abstract justice it does not embody. The abstractions of the moral point of view "risk all the assets of the existing ethical substance".³³ Thus there arises the problem of how "how to make up for this loss of concrete ethical substance, which is initially accepted because of the cognitive advantages attending it".³⁴ The lifeworld must be such as to allow for the application of the norms which are abstractly justified, and to motivate action based upon them.³⁵ This is the problem of mediating morality with ethical life.

Having briefly summarized Habermas's remarks on the tasks and presuppositions of

discourse ethics, I now want to argue that the problem of mediation it leaves us with actually runs much deeper than we are led to suppose. I shall do this by highlighting the following difficulties which face the programme of discourse ethics as outlined above: (1) the reasonableness of the prospect of rational consensus in practical discourse, (2) the problematic status and indeterminate rationality of 'evaluative' statements, (3) the narrowness of the rational content of the good construed as candidature for the just, (4) the problematic identification of the horizon of ethical life with the lifeworld, (5) the tense differentiation between the justification and the application of norms, and (6) difficulties with Habermas's formulation of the problem of mediation. This will force us to reconsider the relation between the reconstructive and hermeneutic moments of critical reflection even on Habermas's *own* terms.

(1) *The reasonableness of the prospect of rational consensus in practical discourse.* Is it reasonable to suppose that participants in discourse about the rightness of norms must hold out the prospect of reaching a consensus? If a positive answer to this question is to be plausible, the content of the norms under discussion would have to be minimized. That this is the case becomes evident as soon as one considers the kind of interests the satisfaction of which would be consequent upon a universally consented-to norm. The threat to consensus, it seems, comes not so much from the diversity, as from the *incommensurability* of interests. There is no reason why diversity (or pluralism) *in itself* should generate dispute, unless the realization of one or more of the diverse interests were to be incompatible with and to challenge the realization of *significant* others. So the prospect of consensus increases in inverse proportion to the scope for incommensurability. But as the scope for incommensurability diminishes, so does the significant content of the norm which

is the object of practical discourse. The problem here is a reminder of Hegel's critique of Kant's principle of universalizability; that the principle is either empty or inconsistent.³⁶ Although Habermas's consequentialist reformulation of the universalizability principle is meant to neutralize this criticism,³⁷ the objection I am putting retains some force. For its thrust is not that universalizable norms must be empty, but rather that norms which hold out the prospect of universal consensus may be so minimal as not to be *worth raising* by concrete participants in practical discourse. Or, if discussion over a norm with such a content is undertaken, argumentation over the universalizability of the norm may not merit the attribute 'rational'. Consider, for instance, John Rawls's fanciful case of the man whose conception of the good life consists in counting blades of grass.³⁸ One can imagine a norm which makes possible this good life being taken up for discussion in practical discourse. It may well pass the procedural test, since the consequences of its general observance for the satisfaction of each person's particular interests could be acceptable to all; but only because it simply doesn't *matter* to them. It could be objected that the example is absurd, since there would be no good reason for adopting such a norm. But this is just the point; that what counts as a good reason for raising a norm for discussion comes *prior* to the procedural test. On the other hand, if the norm is taken up as a constitutive condition of the good life - as *requiring* all to count grass - it might *still* be acceptable if the community were to have such idiosyncratic desire. At this juncture, Habermas would have to follow Rawls's suggestion that the desire be attributed to a kind of collective neurosis. But in making this move, Habermas would be subordinating the task of rational reconstruction to the task of methodically carried-out self-critique. I shall return to this issue of the priority of the two modes of critical reflection enjoined by the theorist of communicative competence below.

To take the basic point I am getting at here further, the *content* of interests must be of a certain recognizable kind if we would call participants in discourse about them rational. That is, participation in practical discourse is conditioned by the possession of a recognizable conception of acceptable generalizable interests, of what is *worth* considering as just. One might want to go on to say that only beings of a certain kind - with a certain identity which discriminates between what is and what is not worth debating - qualify as conversational partners, no matter what the procedure. But this position doesn't follow from the objection I am putting. For we can still say that one only *finds out* what's worth discussing through the process of communication with others. This *dialogical* condition of the validity of normative claims, that they be acceptable to all qua participants in practical discourse, need not be objectionable so long as such discourse is not *analytically* bound to the reaching of a consensual conclusion.³⁹ Unlike the model drawn by Habermas, this view is consistent with the real possibility that participants in practical discourse can disagree, with equally good reason, over the validity of a norm. But more important for our present purposes, it is also consistent with the inclusion of conceptions of the good as proper objects of rational debate over the content of the moral domain.

(2) *The problematic status and indeterminate rationality of evaluative statements.* Indeed, Habermas's reconstruction of the procedural competences of practical rationality leaves statements of what is good - what Habermas calls 'evaluative' statements - in a curious limbo. At one point, he suggests that evaluations or "value-preferences" cannot be rationally debated at all, because they do not hold out the prospect of consensus.⁴⁰ But following from my previous point, it is not the prospect for consensus which makes conversation about norms rational, and even if it were, it is not clear why the prospect

should be any worse for consensus about goods. More typically, however, Habermas recognizes that evaluative as well as moral questions are susceptible to a kind of rational assessment. He suggests that argument over goods is possible but in a different, less strict sense, to practical discourse proper. For the latter, rational consensus is in principle reachable, if only argumentation were conducted for long enough.⁴¹ But this is not the case for argument concerning evaluatives. This kind of argumentation, with its more lenient criteria of rationality, is characteristic of what Habermas calls "therapeutic" and "aesthetic" discourse.⁴²

What status then, do conceptions of the good have? Habermas uncontroversially asserts that they "shape the identities of groups and individuals", which within his architectonic means that they form an intrinsic part of "culture" and "personality".⁴³ Put this way, the good would have to be argumentatively criticizable in terms of the validity claims of objective truth and subjective authenticity.⁴⁴ But since, following the argument of (2.3), conceptions of the good do not apparently fit into either of these categories, to claim this is either to risk distortion concerning the meaning of evaluatives, or if they are to be taken as combinations of these claims, it is to say nothing *distinctive* about them - since *every* speech-act forms a syndrome of the three validity claims. Habermas seems to favour 'therapeutic discourse' as that form of argumentation through which conceptions of the good can be rationally criticized. But how is this possible if there is no *validity claim* which is thematized in therapeutic discourse? A model of *non-criterial* yet rational critical reflection which Habermas requires here is not forthcoming.

It is required because learning processes in the domain of understandings of the good life

cannot be accommodated within Habermas's schema of a logic of development, the rational reconstruction of which is the task of the first of Habermas's two-fold role for philosophical reflection. According to that schema, the capacity to differentiate between the validity claims of truth, normative rightness, and authenticity makes criticism which thematizes each of these claims possible, and thereby conditions cognitive gain - or gains in rationality. But Habermas recognizes that this differentiation also problematizes rational reflection concerning evaluatives or conceptions of the good. He suggests that such reflection integrates what has become differentiated in a harmonious balance.⁴⁵ This is the second of the two tasks he attributes to critical reflection; methodologically carried out self-critique, and interpretation and mediation on the part of the lifeworld. The former, we saw, employs narrative tools for the sake of overcoming particular illusions which block the path to 'self-realization'. But evaluatives, Habermas states, serve to shape or *define* particular identities. Most recently, Habermas has drawn on Taylor's notion of *strong* evaluations to capture the idea of value-preferences which are "not merely contingent dispositions and inclinations", but are rather "inextricably interwoven with each individual's identity".⁴⁶ These evaluations, Habermas states, "both admit and stand in need of justification".⁴⁷ This justification, he continues, comes through "hermeneutic self-clarification", whereby one's life history and process of self-development becomes 'appropriated'. Habermas interprets this appropriation, or "striving for self-realization", in terms of the "*resoluteness* of an individual who has committed himself to an authentic life; the capacity for existential decisions or radical choice of self always operates within the horizon of a life-history, in whose traces the individual can discern who he is and who he would like to become".⁴⁸ Thus 'clinical advice' aims at justified evaluatives, and is addressed to "the resoluteness of the authentic, self-realizing subject".⁴⁹

According to this new position, the rationality of evaluative statements is discernible by way of hermeneutic self-clarification. There are, however, problems in reconciling it with the other claims which issue from the method of rational reconstruction. Take the distinction between who one is and who one would like to become. Each corresponds to two components which are interwoven, according to Habermas, in all evaluatives; the descriptive and the normative. In this case, they correspond to "the descriptive component of the ontogenesis of the ego and the normative component of the ego-ideal".⁵⁰ While Habermas now asserts that "hermeneutically generated self-*description*" (my emphasis) does not issue in "value-neutral self-understanding" - and hence that the descriptive and the normative components of the evaluation are *inseparable* - he is also committed to the view, proposed in defence of the method of rational reconstruction, that the descriptive component of the ontogenesis of the ego is the prerogative of strong *empirical* theories, *independent* of the normative context of processes of development. The two methods of critical reflection are therefore not supplementary, but incompatible. Moreover, if the *description* of who one is is inseparable from its normative context, then it will admit of a *truth-value* which has an inseparable normative force. At this point, Habermas could be taken to mean that the validity-claim raised in the clinical advice is authenticity, not truth. But this cannot be correct, since it is what the self is being authentic *to* which is operative here, something which is independent of, or contingent to, the will to realize it.⁵¹ That is precisely why *resoluteness* becomes an issue; the normative force of the description exercises a pull which transcends my matter of fact capacity to live up to it. But if the validity-claim accorded to evaluatives is truth, the rationally reconstructed differentiation of the validity-claims of truth, normative rightness, and authenticity, must be abandoned.

(3) *The narrowness of the good construed as candidature for the just.*⁵² In his earlier remarks on discourse ethics - where practical reasoning is taken to be the subject of rational reconstruction - Habermas construes conceptions of the good, insofar as they are pertinent to the moral domain, as embryonic justice claims. Habermas's idea here is that conceptions of justice *emerge* out of that horizon of concrete historical traditions and institutions which he calls at various points the lifeworld and the ethical context of life (*Sittlichkeit*). This is what I earlier designated 'ethical totality'. In this context, it refers to the naively accepted background against which communicative actors reach understanding. The claim of these conceptions to rationality, however, and therefore their right to the moral domain, depends upon their successful passage through the procedure of practical discourse. Given that this procedure is what practical rationality means, and that the moral domain is that covered by norms the rationality of which is warranted by virtue of passing the procedural test, conceptions of the good can be of moral significance only insofar as they are candidates for the just.

As will be obvious, this leaves the dimension of goods clustered around the self-regarding virtues of dignity and integrity completely unaccounted for.⁵³ This occlusion can only be explained as the unfortunate consequence of conceiving morality exclusively on the basis of the imperative of social integration through the medium of communicative action. In the terms of the underlying debate which interests us, it is the answer to a question *bound to arise within the discourse of an acultural theory*. Habermas is explicit on the former point;

all moralities coincide in one respect: the same medium, linguistically mediated interaction, is both the reason for the vulnerability of socialized individuals and the key resource they possess to compensate for that

vulnerability. Every morality revolves around equality for respect, solidarity, and the common good. Fundamental ideas like these can be reduced to the relations of symmetry and reciprocity presupposed in communicative action. In other words, the common core of all kinds of morality can be traced back to the reciprocal imputations and shared presuppositions actors make when they seek understanding in everyday situations.⁵⁴

Linguistically mediated interaction, we have seen, is the culture-neutral mechanism by which societies reproduce their symbolic resources. These resources are reproduced rationally when their conviction-carrying power feeds off the illocutionary warrant to justify validity claims when challenged. Habermas gives a plausible account of how the moral intuition of equality for respect can be clarified by way of a reconstruction of the symmetry and reciprocity conditions of communicative, illocutionary binding action-coordination. But it is far from evident how intuitions concerning the moral significance of the self-regarding virtues can be clarified in this way. Put differently, Habermas faces the difficulty of accounting for how conceptions of dignity and integrity are capable of carrying *rational conviction*. It is crucial that Habermas have recourse to such an account, for the following reason. He argues that the transition to modernity can be explained as a process of rationalization, and that the normative content of modernity - self-consciousness, self-determination, and self-realization - can be secured on the basis of its mechanism for rationally reproducing its symbolic resources. What counts as self-realization therefore changes with the transition to modernity. But what *counts* as self-realization is internally related to conceptions of dignity and integrity. Habermas is therefore committed to the claim that *transitions in conviction-carrying conceptions* of self-regarding virtues are subject to rational assessment, *even though* the source of such convictions is not reducible to the procedural conditions of communicative action. But this claim is incompatible with the

view that the normative content of modernity can be secured on the basis of the mechanism for reproducing its rational resources.

An apparent way out of this dilemma is to give conceptions of self-regarding virtues the status of 'clinical intuitions'. This is the path which Habermas takes. It is awkward even on Habermas's own terms, since it is implausible to assess them against a criterial standard of sickness and health. The modification in Habermas's view introduced in response to the previous objection goes some way to obviate this difficulty concerning the self-regarding virtues, since notions of integrity and dignity play a part in 'self-realization', and typically go hand in hand with the virtue of resoluteness. But with this modification, Habermas forfeits the strong claims for rational reconstruction: as the meaning Habermas attributes to clinical intuitions is unpacked, so the relationship between concrete therapeutic reason and formal practical rationality stands in need of reconceptualization.

(4) The problematic identification of the horizon of ethical life with the lifeworld.

Habermas's case for limiting the moral domain to questions of justice, together with his procedural model of practical rationality, turns on the idea that questions about the good can only be addressed from *within* the horizon of the lifeworld. However, the lifeworld as a whole cannot be put into question, since there is no meaningful position to adopt outside of it. Particular aspects of it, can, however, be challenged, once made objects of the hypothetical attitude of participants in discourse. Conceptions of the good, concerning as they do the value of a life as a whole, cannot foster such challenges. The issue then arises of mediating the challenges of practical discourse with the "unquestioningly accepted ideas [and institutions] of the good life" which motivate, give a point to, and concretize action.

There are several difficulties with employing the concept of the lifeworld in this way. First, as a source of the good and identity formations, it is much more complex and internally ambiguous than Habermas's rather homogeneous concept indicates.⁵⁵ Taken as such a source, which can appropriately be called the horizon of ethical life, the lifeworld is not necessarily unproblematic, and not just because as a whole it cannot be treated hypothetically. This is mainly because, as several commentators have observed, the lifeworld is itself the site of conflict between groups who find themselves engaged in a 'struggle for recognition' of the conceptions of the good which give content to their identities.⁵⁶ Second, by levelling all conceptions of the good homogeneously under the same concept, Habermas is unable to articulate the intuition that, since goods vary in their relative preferability and command different degrees of commitment, they must appear as forming (in some sense) a hierarchical order. Third, goods can *lose* their capacity to motivate action. This *is* related to the adoption of a hypothetical attitude towards them, and it does generate a problem of mediation. But, I will argue below, it is a more radical problem than Habermas presents it as.

Conversely, the assumption that moral questions can be rationally discussed, and norms justified, *outside* the horizon of ethical life is also in danger of incoherence. If this background as whole cannot be challenged, then Habermas's claim that practical discourse risks *all* the assets of the ethical life cannot be sustained. Taken individually, Habermas might reply, the particular values can each be put to the problem-solving test. But this move is inconsistent with accepting that the background forms a *web-like* structure, in the sense that the value of its 'nodes' are only genuinely assessable when taken *together*. But more important, even the hypothetical attitude towards particular aspects of the lifeworld,

rather than being tantamount to a transcendence of it, can equally be taken as *peculiar* to the *modern* lifeworld.⁵⁷

Habermas does give some consideration to the preceding objections. Although he fails to make clear in what sense the goods which shape the horizon of ethical life are not just any part of an individual or collective way of life, but rather the means by which particular parts fit together into a worthwhile whole, he recognizes that they too;

transcend *de facto* behaviour. They congeal into historical and biographical syndromes of value orientations through which subjects can distinguish the good life from the reproduction of mere life.⁵⁸

But this recognition fits extremely uneasily with Habermas's claims about practical rationality and the moral domain. For if conceptions of the good are capable of transcending *de facto* behaviour, this means, on Habermas's own terms, that they are capable of motivating *rational* behaviour. But this makes their exclusion from *practical* discourse arbitrary. Admittedly, Habermas has since defused this objection, by distinguishing between what he now calls the pragmatic, ethical, and moral employments of practical reason. But the following difficulty remains. If, by means of conceptions of the good, subjects can distinguish the good life from the mere reproduction of life, such value orientations (the horizon of ethical life) must be distinguished from the 'always already' naively accepted lifeworld. Habermas correctly asserts that the value orientations towards the good congeal into historical and biographical syndromes, and that the kind of rationality appropriate for them is not procedural. It was noted above, when considering Habermas's amended position on the status of 'strong' evaluatives, that they can be *justified*

by hermeneutic self-clarification. But *how* they can be justified relative to each other, and hence what it is by virtue of which they can be ascribed as rational, was left unexplained. Habermas claims that practical rationality can take biographical syndromes as its subject-matter, but how justification of transitions *between* such syndromes - or between strong evaluations - is to proceed, remains unaccounted for.

(5) *Difficulties with the distinction between justification and application.* Habermas's reply to the objection that participants in practical discourse cannot meaningfully transcend the historical context of lifeworld traditions is to make a distinction between justification and application. Participants in practical discourse abstract the values which emerge on the horizon of the lifeworld and put them to the universalizability test. This is the process of justification. But they must then concretize these abstractly justified norms, they must recontextualize them as appropriate moral judgements in the here and now of the lifeworld - the moment of application. In arguing for the necessity of making this hard distinction, Habermas correctly asserts that "no norm contains within itself the rules for its application".⁵⁹ But this assertion cuts both ways, and in such a way as to undermine the conclusion Habermas claims to reach with it. For on the same account, if no norm contains within itself the rules for its application, neither is the application of a rule separable from the rules themselves - a point used by Habermas to justify the initial symmetry and reciprocity conditions of communicative action. The problem is that what a norm *means, qua a rule* for action, is not separable from what counts as appropriate applications of it. It is not just that "moral justifications are pointless unless the decontextualization of the general norms used in justification is compensated for in the process of application", but at a more fundamental level, it's not clear how they could be *meaningful* if this could not

be done. The error, this objection goes, lies in thinking that the application of a meaningful rule is somehow a second step made after its justification. Habermas's reply would be to insist that the possibility of identical ascriptions of meaning presupposes the idealizations of pure communicative action, but this move faces the difficulties I outlined in the previous chapter.⁶⁰

A different kind of objection which Habermas considers is the charge of 'rigorism' made by Hegel of the Kantian approach to ethics. By focussing exclusively on the 'abstract universalism of morally justified judgement', this approach neglects the affective dimension of moral competence and maturity. Habermas's reply is to acknowledge the constitutive role of emotional dispositions and attitudes in the characterization of moral maturity, but only when integrated with the universalizing cognitive operation of the participants in practical discourse. Any adequate description of "the highest stage of morality", he writes, must integrate an ethics of love with an ethics of law and justice. The charge of moral rigorism is only applicable in the absence of such maturity conceptualized as integration, its object is "an impairment of the faculty of judgement".⁶¹

But it is not clear how this highest stage of morality can be articulated through the model of communicative action. Benhabib has argued forcefully that action motivated out of an ethics of love may be oriented about the particular rather than the general neediness of an other.⁶² The other may appear in need of sympathy, encouragement, affection, support. But this could well be in conflict with the requirements of communicative action. Not only may the raising and redeeming of the validity claims of truth, justice, and sincerity be simply irrelevant to the other who appears in such a way, but more radically, such an

appearance requires an orientation which on Habermas's terms can only be conceived as constitutive for *strategic* action. Only by generalizing the other can communicative ethics lay claim to grounding morality. Only such a concept of the other is consistent with a concept of an ideal communication community which transcends the lifeworld as a formal or virtual totality.⁶³

(6) *Difficulties with Habermas's formulation of the problem of mediation.* The impairment of the faculty of judgement is related to the problem of mediation between a universal morality and a particular ethical life. This is, of course, a distinctively modern predicament. It is only possible given the differentiation of the three validity claims and their corresponding value-spheres. The cognitive gain achieved by taking norms as hypothetical universals under the single validity aspect of justice needs to be *compensated for* in the recognition-carrying institutions and character-forming processes - in brief, the socializing medium - of everyday practice. The problem must be seen as part of modernity's need for self-reassurance. As the stringent demands of morality and the facts of ethical life become separated, as the normative content of the pragmatic presuppositions of communicative action are refused embodiment in a rationalized but *colonized* lifeworld, communicative actors become increasingly disposed towards motivational crises, and societies become increasingly vulnerable to legitimation crises. The problems of mediating between morality and ethical life, as Habermas puts it, "centre around the idea of a non-reified everyday communicative practice, a form of life with structures of an undistorted intersubjectivity".⁶⁴

One difficulty which this interpretation of the problem of mediation immediately raises is

that the problem would seem to remain *even without* reification as Habermas theorizes it. It is part of the meaning of lifeworld rationalization that the validity claims and value-spheres separate out. One might thus argue that the faculty of judgement, by virtue of which they are reintegrated in the moment of practical application, is *bound* to be impaired under these conditions. Further, if the exercise of this faculty is to have the force of rationality, and this exercise is impaired under conditions of communicative rationalization, one must also assess this process not as a gain, but as a loss of rationality. For, following from the previous objection, it only makes sense to talk of the rationality of *applied* judgements.

Another difficulty emerges from the conceptualization of motivation and legitimation crises. In line with his deontological approach to the moral domain, Habermas resists making any reference to the good here. As we have seen, Habermas recognizes that issues of the good life "invariably deal with the totality of a particular form of life or the totality of an individual life history" ⁶⁵. He contrasts the foregoing crises with crises in *this* dimension as follows:

A person who questions the forms of life in which his identity has been shaped questions his very existence. The distancing produced by life crises of that kind is of another sort than the distance of a norm-testing participant in discourse from the facticity of existing institutions.⁶⁶

But the two kinds of distance are far from being unrelated to each other. As I noted in my discussion of Gellner, the hypothetical attitude itself *indirectly* undermines the conviction-carrying identities which provide the agent with non-instrumental reasons for

acting. More important, however, is Habermas's reliance on the kind of crisis described in the passage just cited *when dealing with the problem of self-reassurance*. The distance produced by the life-crises of which Habermas speaks is *forced* on practical agents; it has an existential 'unbearability' which impels *change*. In my introduction to chapter four, I noted that Habermas appeals to this model of life crises to refute value-scepticism. The 'ought' which is available to critical reflection upon the communicative mediation of human beings, and which does the work of critical self-reassurance, therefore has a 'clinical' meaning. But this is now revealed as an 'ethical' not a 'moral' ought, and it is scepticism about the distinction between the reproduction of good life and the mere reproduction of life which is challenged. Habermas asserts that the moralizing gaze can directly undermine any particular *Sittlichkeit* insofar as the latter fails to live up to universal standards of justice. As we have seen, the possibility of successful hypothetical problem-solving through communicative action is that the lifeworld stands as an unproblematic background. But this condition is precisely not met in the scenario which generates the need for self-reassurance.

Furthermore, this scenario is difficult to sustain by means of rational reconstruction. For the moment of transition, by which the agent caught up in the causality of fate regains identity, requires the *practical appropriation* of Habermas's *theoretical* perspective on the lifeworld.⁶⁷ Habermas employs the tool of rational reconstruction to ground the rationality of justice claims. He then posits the ideal communication community as the virtual ethical totality through which the abstract participants in practical discourse - wherein "the autonomous will becomes completely determined by reason"⁶⁸ - nourish a universal solidarity. But this is just the "notional republic" which, on Hegel's account, misleadingly gives the claims of justice a *logical priority* over other 'precincts of virtue'. This is why

Hegel introduces the concept of an ethical totality. Habermas takes on this idea to explain the inseparability between the individual and the common good - between autonomy and solidarity. It cannot be explained by rational reconstruction, however, since conceptions of the good, as Habermas acknowledges, do not follow a pre-hermeneutically accessible sequence of stages of competence. This is a fundamental tension between the model of crisis Habermas derives from Hegel, and the validation model for critique which has its basis in the rational reconstructive method of critical reflection.

I have argued in this section that critical reflection in the form of rational reconstruction not only needs to be compensated by hermeutical self-critique, but that the two forms are in tension with each other. Habermas claims that the latter is guided by clinical intuitions, which are subject, in a sense, to rational assessment. Clinical intuitions inform reflection on 'ethical life', which are distinguished from valid practical judgements in virtue of their lack of universalizability. The Kantian conception of practical reason as 'unconditioned' is at once claimed and denied; conceptions of the good can motivate rational behaviour, and evaluatives are recognized as having normative assertoric force, while communicative competence differentiates between truth-value and norm-value, and opposes rationally motivated agreement from empirically motivated assent. It seems as if a rethink is required of the division between rational reconstruction and interpretative mediation, and correspondingly of the roles that therapeutic and practical reasoning play in thinking about the problem of self-reassurance.

(5.4) Therapeutic and Practical Reason Reconsidered

In the terms Habermas's critical project sets for itself, the problem of modernity's

self-reassurance falls under two aspects. First, there is the problem of how modernity can generate its norms out of itself. The task here is of reconstructing the normative content of modernity out of the presuppositions of the vehicle for reproducing its symbolic resources - communicative action. The critical standard is a formal notion of intact intersubjectivity; the procedure of rational will-formation for participants in an ideal communication community. Modernity's normative content - self-consciousness, self-determination and self-realization - owes its claim to rationality to the unavoidability of the idealizing presuppositions of communicative action. By the same token, this rationality can only claim to be procedural. So, for instance, the space for self-determination which *structures* a form of life can be criticized according to the ideal standard of the universal norms generated through the procedure of practical reason. The lifeworld is *simultaneously* the transcendental site at which ideally rationally motivated speakers and hearers meet, and the historically contingent context and resource of actual communicative actors. Second, there is the problem of a lifeworld which is one-sidedly rationalized. But the lifeworld can only be one-sidedly rationalized *as a whole*. There is no validity claim by virtue of which critique is possible here. The standard of intact intersubjectivity can only come from the clinical intuitions (about 'sickness' and 'health') of the critic. This division of critical labour thus corresponds to the ambiguity in the standard of intact intersubjectivity I identified in (5.1). Insofar as it stands for the dialogue situation of pure communicative action, it does not cover the problems of mediation which remain to threaten modernity's self-reassurance. These problems can only be tackled by appeal to the standard of a balanced lifeworld, which cannot be derived from rational reconstructions or the exercise of practical (as distinguished from therapeutic) rationality. Even worse for Habermas is the prospect of *conflict* between the standards appealed to by participants in practical and

therapeutic discourse. For although he considers the modern lifeworld to be one-sidedly rationalized in its cognitive/instrumental aspect, it is equally possible that there may also be a one-sided development of its moral/legal dimension. Given Habermas's model, such a situation would still satisfy the requirements of practical reason. This is another reason for thinking that practical and therapeutic reason are not only divided, but potentially in conflict.

The relationship between practical and therapeutic reason, and correspondingly the status of clinical intuitions, thus calls for reconsideration. In fact, Habermas has repeatedly acknowledged the difficulties which the notion of clinical intuitions brings. When asked about the appropriateness of a standard of sickness and health for critique, Habermas replies revealingly that for intuitions concerning the value of a form of life as a whole;

we apply yardsticks which are valid in the first instance in the context of our culture or plausible in the context of our tradition... So far I have no idea of how the universal core of those merely clinical intuitions - if indeed they have one at all - can be theoretically grasped.⁶⁹

The point is decisive; if there is one unresolvable problem which threatens Habermas's project for critique with exhaustion, this is probably it. The problem arises from the following theoretical commitments. First, Habermas is committed to the view that normative universality is only graspable from a position of impartiality. Second, he claims that impartiality requires a transcendence of the culture in the context of which validity is provided in the first instance. He thus needs to show how the universal core of clinical intuitions can be theoretically grasped *independent* of particular cultural 'yardsticks'. But this is an impossible task, since clinical intuitions can never have application outside of

particular cultural contexts. Habermas seeks a criterion of sickness of health which is universal in virtue of its impartiality between cultures, in the same way as the criterion of justice is. But sickness and health (in the present context), Habermas also claims, are inseparable from need-interpretation, and not reducible to medical or biological invariables. That Habermas requires a yardstick which has a universal core in virtue of being culture neutral betrays his commitment to an acultural theory, *yet the conceptual resources of his acultural theory are unable to yield just what needs to be theoretically grasped.*

In order to be theoretically graspable, the claim to universality of clinical intuitions would need to be conceptualized non-criterially; their validity would have to be construed in a way which is bound to particular cultural contents, but not merely in the *first* instance. As we have seen, in his most recent writings Habermas turns to the appropriative understanding of hermeneutically generated self-clarification for a model for how *changes* in clinical intuitions are rationally justifiable. Insofar as they are rationally justifiable, they have some claim to universality. But there is really little new in this position, as a reminder of Habermas's model of self-critique implicit in his reading of Freud, relieved of the burden of representationalist assumptions exposed by Bernstein's anti-naturalistic objections, might serve to make clear.

As we saw in chapter three, the central themes of that reading were: that a psychoanalysis takes the form of an emancipatory narrative; that the undertaking of the narrative involves taking moral responsibility for the sickness; and that the emancipatory movement (the cure) follows the dynamic of the Hegelian causality of fate. As we saw, the task of analysis is to reconstruct a disturbed self-formative process. The meaning of symptoms, in the here

and now, is incomprehensible to the subject. They must be rendered intelligible by a depth-hermeneutics, an interpretation which reflects back upon the original scene of disturbance. But since this interpretative reconstruction covers a necessarily temporal process of self-formation, it must have a narrative form. Only in this way can the meaning of repressed past events be recovered as *interruptions* in the self-formative process. But in therapy, the self-narration in virtue of which the self-formative process is understood, changes the narrating self. But what is it for this change to take place in the therapy? It involves the adoption of a new framework of self-understanding, one within which different affectively charged attitudes can be taken towards oneself. But this is only possible if the analysand has a will to change herself - what Habermas calls a passion for critique. Thereby moral responsibility is acknowledged for the feelings and actions which are the source of her misery. But to adopt moral responsibility here is to recognize oneself for who one really is, as revealed retrospectively in the act of self-narration. Who the self is is internally related to this moral responsibility, and so to the passion for critique. It is something which must interpret and is capable of reinterpreting its needs, according to narratives which can retrospectively distinguish between a more or less fulfilled life. To act on the former, to continue an interrupted self-formative process, is to continue a life which could retrospectively be seen as *better* as a whole. And this is just what it means to take moral responsibility for the self-formative life process.

The passion for critique, then, requires a framework within which discernments can be made about the worth of competing self-formations or life-narratives. Only with conceptions of the good *internalized into identity* are such discernments possible, in that they provide the motivating power for resolving *crises of identity*. Therapeutic success

returns this power by enabling self-repossession through a framework of self-understanding. But there is no presupposition of an ideal (in the Kantian sense) self to be aimed at. 'Clinical' judgements of what counts as a distortion, in this case, are only available retrospectively within an interpretive framework through which the subject narrates itself. It is not available to the logic of development which the idea of a pure communicative mediation of subjects represents. Not only does therapeutic critique not need it to be possible to transcend such frameworks to a virtual totality of communicative action, but successful therapy is inconsistent with the transcendence which is required of the critic of systematically distorted communication. Since the success of therapeutic critique is presupposed by Habermas's project of critical self-reassurance, the latter is inconsistent on its own terms.

During the same interview in which Habermas acknowledged the ungraspable rational core of therapeutic discourse, he makes what is tantamount to a resigned acceptance of the irresistible conclusion that clinical intuitions are theoretically acceptable only as premonitions of the good life. In reply to objections that the psychoanalytical model of critique presumes a privileged epistemic and evaluative standard on the part of the theorist, Habermas states that the emancipatory psychoanalytical narrative does not define how the life of the individual must continue, only that it return the subject to the dignity of *homo sapiens*.⁷⁰ But of course, what is meant by human dignity is empty outside of an evaluative framework which articulates conceptions of what makes for a good human life. Habermas's theory of crisis and his conception of critique carry force *given* a modern evaluative framework. Taken together, the preceding arguments thus suggest that the critique of forms of life as a whole is misconstrued as therapeutic discourse for the same reason that practical

discourse is misconstrued as a procedure. These misconstruals, I have proposed, can be corrected only upon recognition of the irreducibly substantive nature of practical rationality. Once recognized, the excessively *immodest* burden carried by clinical intuitions, and the excessively *modest* burden carried by procedural practical rationality, can be relieved. Critique oversteps itself when set against a standard of sickness and health, it cuts itself short when set against the standard of procedural unity: the standard of intact intersubjectivity is either too strong or too weak.

(5.5) Conclusion

Habermas acknowledges that the rational reconstruction of the moral point of view along Kantian lines is not without unhappy consequence. By separating the right from the good, and by construing the former as the basic moral phenomenon, it abstracts the agent from the requisite motivations to act morally. By privileging the moral competence which manifests itself at the level of the justification of maxims of action by reference to a principle of universalizability, it abstracts the agent from the particular lived situations in which norms are applied and competence tested. And by according priority to general principles of morality over particular contexts of ethical life, it lends itself to atomistic and contractarian conceptions of the person and society. But he insists that there can be no simple cancelling out of the dilemma between form and content. "The neo-Aristotelian way out of this dilemma", he writes,

is to argue that practical reason should forswear its universalistic intent in favour of a more contextual faculty of judgement. Since judgement always moves within the ambit of a more or less accepted form of life, it finds support in an evaluative context that engenders *continuity* among questions of motivation, empirical issues, evaluative issues, and normative issues.⁷¹

Against this, discourse ethics refuses to go back prior to Kantian thought, insisting that the idea of impartial application is preferable to the idea of prudential judgement as a conception of practical reasoning. Habermas gives a two-fold reason for its preferability. First, the neo-Aristotelian approach is said to be encumbered with "metaphysical premises" incompatible with the evolution of the learning process thought to characterize the transition to modernity.⁷² Second, once this metaphysical basis for critique is renounced, Habermas supposes that neo-Aristotelianism assumes a relativistic and thereby essentially *conservative* character; it seems to offer no rationally forceful protection against prejudice and parochialism.⁷³

The metaphysical premise of the neo-Aristotelian approach, it seems, lies in its presumption of *continuity* between motivational, empirical, and normative issues. For the neo-Aristotelian, how the objective world is (an empirical, quasi-motivational issue) sets the standard for how to act (a normative, quasi-motivational issue), and this provides the basis of the objectivity of a moral order. As a neo-Kantian, Habermas insists that an objective moral order can be grounded only on the basis of universalizable norms of action discontinuous with 'empirical issues', since the empirical issue of how the objective world is can offer no reason for moral action. Only the latter position is consistent with the evolution of learning processes in modernity; the learnt capacity to differentiate and thematize truth-, right-, and authenticity-aspects of validity claims. But recall now the status of clinical intuitions. The coherence of Habermas's concept of clinical intuitions, I argued, depends upon them being construed firstly as premonitions of the good life, and secondly as truth-evaluable *in their action-orienting aspect*.⁷⁴ But this is just the 'metaphysical premise' of the neo-Aristotelian. So far as therapeutic reason goes, then,

Habermas is *already one step down* the neo-Aristotelian path, since the presuppositions of therapeutic reason are *no less metaphysical* than those proposed by the neo-Aristotelian, thus far defined.

This conclusion is corroborated by an objection Habermas puts to Bernard Williams, whom he places in the neo-Aristotelian camp. "How truthfulness to an existing self or society is to be combined with reflection, self-understanding, and criticism", Habermas cites Williams as asserting, "has to be answered through reflective *living*".⁷⁵ He then objects that "Williams is compelled to attribute to practical reason a form of rationality which goes beyond sheer common sense but whose difference from scientific rationality remains to be determined".⁷⁶ But here again, driven by the need to defend his neo-Kantian rational reconstructions, Habermas neglects the third path which he himself has opened up in his conception of theory-mediated, autobiographically informed reflective living, which has a claim to rationality *neither* reducible to common sense *nor* to scientific rationality; namely, the claim to therapeutic rationality of justified clinical intuitions.

If Habermas already has one foot on the 'neo-Aristotelian' path, perhaps the reason he does not take it lies in the second of his objections; that neo-Aristotelianism displays a bias towards conservatism, and thus provides an inappropriate philosophical basis for critical reflection. "In modern societies", he writes,

we encounter a pluralism of individual life-styles and collective forms of life and a corresponding multiplicity of ideas of the good life. As a consequence we must give up one of two things: the claim of classical philosophy to be able to place competing ways of life in a hierarchy and establish at its acme one privileged way of life over against all others; or the modern principle of

tolerance according to which one view of life is as good as any other, or at least has equal right to exist and be recognized.⁷⁷

And he continues;

if we wish to remain faithful to the Aristotelian conviction that moral judgement is bound to the ethos of a particular place, we must be prepared to renounce the emancipatory potential of moral universalism and deny so much as the possibility of subjecting the violence inherent in social conditions characterized by latent exploitation and repression to an unstinting moral critique. For only the posttraditional level of moral judgement liberates us from the structural constraints of familiar discourses and established practices.⁷⁸

Certainly, if the neo-Aristotelian position entails that competing ways of life be dogmatically placed in a hierarchy with one particular way of life unassailably imposed at its highest point, if it really does rule out the very possibility an unstinting moral critique of latent exploitation and repression, it is unfit as a philosophical basis for critical reflection. But the alternatives presented in these passages do not exhaust the possibilities for critical reflection, even on Habermas's own terms. In the previous section, we saw how Habermas appeals to clinical intuitions in order to return the subject to the 'dignity' of homo sapiens, a claim which implies the privileging of at least one way of life over others precisely insofar as they compete with each other. Furthermore, the critique of "the violence inherent in social conditions characterized by latent exploitation and repression", is *not* 'moral' strictly speaking, but *depth hermeneutic*. And the conditions of acceptability of such critique, as well as the emancipation it issues, we have seen, are those suitable to clinical intuitions. Once again, in order to defend his neo-Kantianism, Habermas ushers from view the very theoretical resources which he otherwise makes available for critical reflection. Bewitched

by the idea of critique as requiring an ideal standard, practical reasoning can be misconstrued by Habermas as a procedure which can be rationally reconstructed. If Habermas were to dispense with this idea, ways of life might be amenable to *rational* comparison, not in terms of an evolutionary scale of problem-solving, but through competing narrative discourses of 'world-disclosure' concerning the constitution of worthwhile ways of life. This would mean that criticism of an alienated or damaged form of life as whole - the kind of criticism germane to the problem of modernity's critical self-reassurance - could claim to be both practical and rational.

Taylor attempts to place this kind of moral critique, which he takes to be suitable to the fundamental moral predicaments of modernity, on a philosophically sound basis. Habermas acknowledges that Taylor, though a neo-Aristotelian, seeks to open up a space for critique which buffers his position from the charge of conservatism. But while not ruling out the possibility of unstinting moral critique, Taylor's "universalistic ethics of the good that appeals to supreme goods transcending all particular forms of life", Habermas asserts, "are grounded in cosmological and religious worldviews that are even more difficult to reconcile with postmetaphysical thought than the teleological worldview of Aristotle".⁷⁹ In the next three chapters, it should become clear that Habermas's criticism betrays quite a radical misunderstanding of Taylor's position.

CHAPTER SIX: TAYLOR'S CONCEPTION OF CRISIS

(6.0) Introduction

Habermas's model of critical analysis is supposed to be applicable wherever societies must reproduce themselves through linguistic interaction. Since this requirement holds for all societies, whatever the substantive content of the norms which bind them together, Habermas's theory is in a certain sense culturally non-specific. Societies are rational to the degree that the norms which bind them together are accountable to the constraints imposed by the procedure of communicative action, which is the degree to which these norms transcend their culture-specificity, or are universalizable. The rationality of cultures can be assessed in terms of the degree to which they approach being *structured* by communicative action. But the normative constraints imposed by the structure itself are not specific to any particular culture. The precise nature of the acultural standpoint of Habermas's theory of modernity roughly sketched in (1.1) should now be evident. While Taylor correctly identifies Habermas's insistence on the 'immanent logic' of the modern differentiation of the three 'validity spheres' as a manifestation of the acultural leanings of his theory, it is Habermas's focus on the general capacity for linguistically mediated interaction which makes it acultural at an even more fundamental level. However, we are also now in position to modify our initial characterization. For as I argued in the latter sections of the previous chapter, the role which a concept of crisis plays in Habermas's conception of critique renders it less acultural than my preliminary outline made it seem. Not only is there a tension between Habermas's (acultural) rational reconstructions and his (cultural) model of narratively carried out self-critique, but, I have argued, the latter must enjoy priority insofar as his theory addresses the question of modernity's self-reassurance.

But I have not yet considered how the question of modernity's self-reassurance can be reformulated, nor have I assessed the prospects for successfully tackling it, once a more thoroughgoing cultural-theoretic approach to modernity of the kind proposed by Taylor is adopted. This is the task I shall undertake in this and the following chapter.

In this chapter, I will (6.1) offer a clarification of what a cultural theory of modernity, of the kind advocated by Taylor, is supposed to be *about*. Here I will outline the core concepts around which Taylor's proposed theory of modernity is to be organized. I will then be in a position to reformulate the problem of self-reassurance, by means of the conceptual resources available to such a theory. The keys to this resource are the internally related concepts of 'identity' and 'the good', which Taylor bridges by way of a distinctive concept of 'strong evaluations'. My aim here is to provide enough of an exposition of these basic categories to give us a handle on how the problem of self-reassurance raises itself when thought through them. I then turn (6.2) to how Taylor does indeed raise it. Although, as I mentioned in (1.1), Taylor's official view is that his suggested way of formulating the problem of self-reassurance represents only one possible means of addressing it, my claim will be that this is an excessively modest self-presentation on Taylor's behalf. For Taylor argues implicitly that the problem as raised by the cultural theory is a question which *must* be formulated, and that a cultural theory as Taylor defines one is *indispensable* for a critical self-understanding of modernity. As I have attempted to show is the case for the grounding of Habermas's theoretical approach, this is a claim which appeals to a particular conception of crisis. I look at this conception, and interpret the nature of the claims associated with it, in (6.3). After raising some initial objections to this claim - that human agency is necessarily articulable by a framework of strong evaluations

(or the kind of contrasts which cultural theories investigate), that doing without strong evaluations is impossible for human beings - I turn in (6.4) to objections which can be reconstructed in the light of an alternative perspective on the self. Specifically, I do this by considering how well Taylor's view stands up to the challenge of a rival 'pragmatist' conception of the self - one that allegedly does do away with strong evaluations - put forward by Richard Rorty. I then raise some problems which, even if Taylor's thesis concerning the inescapability of strong evaluations were sound, still face his cultural theory. I conclude the chapter (6.5) by drawing some parallels between this and Habermas's theory, for the sake of contextualizing the need for Taylor to overcome the aforementioned problems.

(6.1) Persons, Strong Evaluations, and the Good

In characterizing 'cultural' theories of modernity by way of their contrastive focus on modern conceptions of personhood and the good, Taylor is trading on what he claims are internally related concepts of 'the person', 'the self', 'identity', and 'the good'.¹ In this and the following two sections, I shall offer a largely expository account of what I take to be the core of these claims. I will be particularly concerned with clarifying the sense in which an unseverable connection between human identity and the good is urged by Taylor, and how this sense can be elucidated from the kind of argument which supports it. To show that the concept of a person is in a sense unintelligible independent of the lived capacity to draw strong qualitative distinctions, and that these distinctions form a framework which must orient human agents towards the good, is the burden of his phenomenological account of identity crises. But that account is supported by a cluster of concepts - like 'personhood', 'strong evaluation', and 'the good' - which organize his favoured 'cultural'

theory of modernity, and through which its need for self-reassurance can best be articulated. I first want to look at what Taylor thinks we mean by these concepts, and then to how the problem of self-reassurance can be formulated in their terms, before turning to his phenomenological argument for showing that we must articulate our understanding of them in that way.

If there is a single core idea around which Taylor's theory is organized, it is that the concept of a person is essentially that of a being for whom things *matter*.² The distinctive themes of Taylor's proposed cultural theory of modernity can, I think, be traced back to this basic insight. First, if it is true that a person is a being for whom things matter, it follows that things necessarily matter to a person the *more or less*. Something always matters to me rather than something else, and if I could not discriminate between the two, if I could not contrast one against the other, I would have no concept of things mattering to me at all. Put another way, if everything mattered the same, if anything mattered, nothing would.³ What matters makes a *difference*.⁴ Take, for instance, it mattering to me that I keep faith to a principle, or to a person, or to class of persons. It mattering to me is intelligible against a background of contrast; what a life would be like which betrayed the principle, the person, or the class of persons, in failing to keep faith. Keeping faith matters on the understanding I have of what it contrasts with - in this case betrayal, or alternatively not caring less either way. Or take it mattering to me that I be a person of generous spirit. Again this presupposes that I understand what it is to be mean-spirited, and that I care one way or the other about which of the contrasts are true of me. The degree to which I care will of course depend on the particular object of it, and there may be more than one contrastive term through which I understand *that* it matters to me. To take the latter point

first, it might matter to me that I have an open rather than a closed mind, but the contrast to 'open-minded' may also be made by reference to such terms as 'committed', 'resolute', or 'critical'. And of course I might change my mind about what matters in the light of these alternative contrasts. But this does not detract from the point that grasping the concept of a person involves a grasp on the concept of mattering as articulated by *some* set of contrastive alternatives. To turn to the former point, it might matter to me that I resolutely keep faith with a principle more than keeping an open my mind to the ambiguities or rigidities of principles in general, even though both resoluteness and open-mindedness matter to me. And I might change my mind on this too. It is crucial to Taylor's position that these changes of mind can stake a claim to rationality, the nature of which will be examined in the next chapter.

The concept of a person to which a cultural theory has recourse, then, is that of a being for whom things matter, a being which cares about things.⁵ What matters to me is by definition not something I 'couldn't care less' about, I cannot be *indifferent* to it. But once we accept that things matter the more or less, Taylor wants to urge, we are committed to acknowledging that there will be things which matter most. What fundamentally matters is signified, so to speak, by 'bottom-line' oppositions between the worthwhile and the worthless, the significant and the trivial, the fulfilling and the empty, which define what Taylor calls the 'incomparably higher'. The worthwhile, the significant, the fulfilling, really matters; the worthless, the trivial, the empty, either doesn't or only appears to do so. And a person's being matters to the degree that it is actually or potentially worthwhile, significant, or fulfilling. As Taylor conceives it, a person is a being whose agency makes sense in virtue of such distinctions. But it is now also clear that distinctions of this kind

are also *evaluations*, and that what matters to human beings is that their lives be oriented to the good. At the most general or fundamental level, what makes the life of a person or group of persons worthwhile, significant, and fulfilling, is that which defines the good life for that individual or group.

Taylor identifies three different strata to the good life which correspond to three analytically separable axioms of moral intuition: that a course of life can either be endowed with or be lacking of meaning; that it can either succeed or fail to live up to ideals such as dignity; and that a form of life can either respect or flout rights to freedom, self-determination, a homeland, bodily integrity, basic material well-being, and the like.⁶ In each case, so Taylor contends, the good must be defined contrastively, and definitions of goods taken together make up a 'framework' which furnishes a person with an orientation for acting for the best, or living to the full. For Taylor then, a person as a being for whom things matter is a being which exists against a 'background picture' or 'framework' of qualitative evaluative contrasts. Human agency requires a 'horizon' which discriminates between ways of life or modes of being which matter the more or less, which are of greater or lesser significance for *human* being.

The force of this last claim turns on the *kind* of evaluation which Taylor thinks is at stake in making such qualitative contrasts. Taylor makes a distinction between two different kinds of evaluation, which he calls 'strong' and 'weak'. At issue in a weak evaluation is the weighing up of *de facto* preferences. For instance, faced with a decision to have coffee or tea, I weigh up which I fancy most, and go for coffee. There is a sense in which it matters to me that I have coffee rather than the tea, because I just happen to desire it, and

would rather have that desire satisfied than another I might happen to have. By contrast to a weak evaluation, for which it is sufficient that something (in this case the coffee) be de facto desired, a strong evaluation involves qualitative distinctions concerning the worth of the motivation. Strong evaluations invoke normative judgements concerning the relative worth of alternative desires. These stand independently of one's de facto desires, and offer standards by which their value can be judged. To keep with our example, there are circumstances where I might strongly evaluate between drinking coffee or tea, and this would change the *sense* in which it mattered to me which I were to have. For example. I might crave for the taste of the coffee but strongly evaluate in favour of tea because the coffee has been produced exploitatively. In this case, I might think that choosing the coffee would make me a worse person; a thought which couldn't occur to me if I were merely weakly evaluating in favour of the coffee. Or I may hate the taste of coffee but go for it because it is an export of an ideologically sound or economically needy country. The point here is that in both cases it is my *stand* which matters, not my de facto preference for coffee or tea. And insofar as I evaluate by appeal to a standard which is not contingent upon my de facto desires, I evaluate strongly. In Taylor's terms, this stand would be part of my identity. For this reason, only strong evaluations are relevant for what is distinctive about human agency, as Taylor is portraying it.⁷

In the example just given, I might well be motivated to take my stand through perception of the good of benevolence - through an understanding of benevolence as part of the good life. Benevolence is thus an example of what Taylor calls 'life goods'. But *amongst* these life goods, there will be some which matter more than others. In other words, besides distinguishing between strong and weak evaluations, we also need to discriminate between

the *relative strength* with which different strong evaluations are individually held or institutionally embodied. Those goods which correspond to the top of the hierarchy of strong evaluations Taylor calls 'hypergoods'. Hypergoods are "goods which not only are incomparably more important than others but provide the standpoint from which these must be weighed, judged and decided about".⁸ They are 'higher-order' goods, in an analogous sense to which goods are objects of 'second-order' desires or (strong) evaluations.⁹ To say that hypergoods are 'incomparably' more important than other goods is also to say that any one is 'incommensurable' with any other - that they cannot be measured against each other - since the measure of what is important is itself defined by the specific hypergood. Hypergoods therefore stand as rivals for the moral allegiance of persons and societies.

But in order for allegiance to hypergoods to be possible, Taylor thinks, we must suppose there to be some reality which stands behind or constitutes them as good. In distinction to life goods then, there are also what Taylor calls 'constitutive goods'. For a hypergood to be the object of allegiance there must be something constitutive of it which motivates that allegiance. Taylor defines a constitutive good as "a something the love of which empowers us to do and be good".¹⁰ It is that by reference and appeal to which one is *inspired* towards that which has a call on me. Because constitutive goods have this empowering quality, they function as what Taylor calls 'moral sources'. A moral source is "something the contemplation, respect, or love of which enables us to get closer to what is good".¹¹ Taylor mentions God, Plato's Form of the Good, and the power of rational agency articulated by Kant, as examples of such sources.¹² It is contemplation of the Form, love of God, or respect for the rational agent which puts the person in contact with what is most important, worthwhile, or most incumbent upon persons. And such contemplation, love, or respect,

is the manner in which the awe and allegiance which is proper to the constitutive good manifests itself. But there is another, even more fundamental way, in which this reality is disclosed.

For something to matter for a person is for it to be invested with significance, and the investment of significance is something which is disclosed by language. A person, as a being for whom things matter, understands the significance of things by interpreting them through language. Since the particular identity of a person consists in what particularly matters for that person, and since what matters or is of significance to a person is disclosed by interpretation through language, persons must be considered as self-interpreting beings.¹³ The articulation of more or less significance is what, on Taylor's account, self-interpretations do. So for a person to have an identity is to relate to an interpreted world which matters the more or less. Although it is perhaps more usual to mean by 'identity' a set of physical properties which uniquely individuates an object (rather than a source of interpretation), since the peculiarity of being a person involves interpreting who one is, we can usefully talk of different ways of being a person according to different conceptions of 'what it is to be human' available for self-interpretation. To get clear on these conceptions is to articulate and find one's place within the 'background picture' or 'frameworks' within which one finds one's identity. And insofar as historical variations can be discerned in the organizing concepts or frameworks of self-interpretation, we can make such very general distinctions as between 'modern' and other identities.

Given that a person is partly constituted by a language of self-interpretations, it is short step to the thought that "One is a self only among other selves".¹⁴ What matters to a person,

since it is articulated in language, is not just an expression of my being, but of my identity as necessarily a participant in a collective practice which is expressed and supported by that language. The object of significance, disclosed by language, is meaningful (at least in the first instance) for the 'us' of a language community. One is a self only among other selves because selfhood only emerges in the medium of 'webs of interlocution'.¹⁵ A person can come to self-understanding and self-definition only in relation to other interlocutors, not only because the resources for his or her self-interpretations are publicly shared meanings, but also because I come to grasp these meanings through relating or interacting with others. The need for this interaction in coming to grasp meanings Taylor calls our 'defining situation' or 'transcendental condition'.¹⁶ But this condition can become problematic. When it does, there issues the need for self-reassurance.

(6.2) The Problem of Self-Reassurance Reformulated

The problem of modernity's self-reassurance, seen in the light of the conceptual resources available to a cultural theory, refers to the sustainability of a historically and culturally specific identity. The modern identity includes those conceptions of the good life and notions of what it is to be human which are dominant in modern societies, and which are expressed and carried in their characteristic languages and institutions. The problem of modernity's self-reassurance issues from the 'fateful tendency' of the modern identity to undermine itself. Taylor puts the problem like this;

We live in a society whose practices embody a certain notion of identity, and the human good. This must be ours or we cannot give this society our allegiance; we are alienated from it. At the same time, we rely to a great extent on these practices to maintain this sense of identity. If these practices which supposedly embody the modern identity can be shown to lead in fact to some such failure to achieve it...perhaps our faith in the conception of the modern identity is shaken as well. We turn to other models.¹⁷

This is Taylor's model for understanding the legitimation crises to which modern societies are disposed. His supposition is that legitimation crises arise when social and political institutions fail to embody the frameworks of strong evaluation which legitimate them. The allegiance of the participants in the modern identity can be lost either because they find they can no longer take their stand within the dominant framework of qualitative contrasts, or because the institutions which lay claim to carrying this identity fail to do so. They either cannot orient themselves within the moral space defined by the modern horizon of strong evaluation, or else they detect such a radical gap between these evaluations and the practices which are supposed to express them that the framework itself loses its hold. Of course on Taylor's account, the identity of a practice just means the framework of goods it presupposes and reproduces. And as this suggests, Taylor portrays the 'fateful tendency' of the modern identity as arising in large measure from a conflict *between* the hypergoods of modernity. It is as partisans of conflicting hypergoods that, according to Taylor's theoretical approach, the exponents of 'Enlightenment' against 'Romanticism' and vice versa should be understood. This is one major source of the problem of self-reassurance, since we all, to some extent, live under the pull of these two great "frontiers" of the modern identity.

Although I will not consider either 'stand' in any detail, I still hope to be able to show in summary what, as approached by Taylor's 'cultural' theory, is at stake in this conflict. But this conflict, which issues in the critique of instrumentalism, is not the only source of tension for the modern identity. The 'disengaged' stance, which Taylor claims is espoused by defenders of 'Radical Enlightenment', although central to the modern identity, simultaneously undermines it in other ways. Most crucially, under the guise of the

hypergoods of benevolence, ordinary life, and an ethic of cognition, it disavows the *constitutive* goods on which the hypergoods depend. Moral sources are thus lost from its purview; constitutive goods are dismissed as illusion. But since moral sources are required for motivating persons towards hypergoods, the modern identity, insofar as it is forged around an ideal of disengagement, tends to frustrate its own articulation. And this gives the edge to Taylor's own conception of philosophical critique as a contribution towards a cultural theory of modernity; to articulate in a language of perspicuous contrast the sources of our and other identities. But this is a hasty anticipation of how such a theory might answer the problem of self-reassurance. Let us first look a bit more closely at how the theory asks it.

According to Taylor, the modern identity has in large part been forged around two competing hypergoods, which he identifies as rational disengagement and expressive unity with nature. Taylor describes how for the pre-modern the source of significance for human life lies in the cosmological order. To find meaning in life, to articulate his or her ideals, and to locate the source of his or her obligations, the pre-modern looks *outwards* to the order of things in which he or she finds his or her place. Taylor takes Platonism as a paradigm articulation of this kind of identity. One makes contact with the good, the true, the worthwhile, and the just, by turning one's gaze to the order of the cosmos; the (real) world of Forms. By contrast, the horizon of identity for the modern lies 'within'. But this turn inward for the purpose of locating *moral sources*, through which the modern finds his or her 'self', can be made in two quite different and incompatible ways, which correspond to the two chief rivals amongst modern hypergoods.

On the one hand, I can *disengage* myself from the thoughts, feelings, and desires which happen to constitute me as a natural being. Although this is a point on which Taylor is not always consistent, it is not the substance of these inner phenomena which defines my 'self' as a modern, but how I find a normative orientation in relation to them.¹⁸ Taking a disengaged stance is one such way. In adopting a stance of disengagement, I exercise a capacity the strong evaluation of which is a distinctive facet of the modern identity: That is, it is my ability to stand back from my thoughts and purposes, to resist being fooled by the force of habit and tradition, and to reshape them in a rational order according to some neutral methodical procedure, which confirms my dignity as a human being: It is that in virtue of which I am the fit object of attitudes of admiration rather than contempt. I merit admiration because of my adherence to the ideal of 'self-responsible reason' - to courageously affirm my responsibility to my own individually, rationally thought-out convictions, and freedom of belief in the face of whatever authority.¹⁹ The self-mastery which my disengaged, self-disciplined stance makes possible is directed towards my tendency towards prejudice and illusion. This also allows me to endow what is outside me (the so-called 'external world') with significance by mastering *it*. The external world has already lost its intrinsic significance through the modern turn inwards (strictly speaking it only comes into being *as* an external world with this turn), but it reappears now as of *instrumental* value to the disengaged human subject. Disengagement means the adoption of an instrumental stance towards inner and outer nature by means of which it can be controlled with maximum efficiency. The characteristically human, the source of human excellence, is in this respect "the capacity to dominate, and not be dominated by things".²⁰ Taylor portrays the dominant practices of modernity's technological civilization as embodiments of this strong evaluation.²¹

On the other hand, and in reaction to the disengaged stance, there are moderns who identify with those inner feelings and impulses which are unsullied by the distortions and mutilations of self-control and the objectification of self which disengaged control requires. According to this framework, what is admirable about human existence is the capacity for spontaneity, playfulness, and self-exploration, rather than for calculated and efficient instrumental thought and action. Genuine human dignity is thought to lie in an openness to the voice of nature within. To be heedless to this voice, to be insensitive to it, is to be deaf to what really matters in life. And to act on motivations which fail to attune with this inner nature is to merit an attitude of either indifference if this is of no consequence, or of contempt if it results in the manipulation of the inner nature of others or the destruction of the natural environment. Either way, on this horizon disengagement represents self-mutilation rather than self-fulfilment. According to this facet of the modern identity, the turn inward is of no good if disengaged - for nature as it calls us within is not accessible to a rational, objectifying gaze. Rather, it requires the exercise of a creative imagination, and in its exercise human excellence is made manifest. This stance of self-exploration rules out disengagement *a priori*. The instrumental stance presupposes a knowledge of the object to be controlled, precisely what the exploratory stance cannot assume. To resist this presumption - to be unprepared to dominate things and to create a language which articulates an alternative relation to one of domination - is to be best disposed as a human agent. Again, Taylor portrays much of the protest against technological civilization as an expression of these strong evaluations.²²

What needs to be emphasized here is that *both* the Enlightenment horizon of disengagement and the Romantic horizon of self-expression are portrayed as internal to the modern identity

by Taylor. Both share a sense of the self as 'inner', and they rely on each other for the sense in which their conception of the self is to be judged as superior. This last claim is crucial, for Taylor insists on avoiding the kind of manichean oppositions which have bedevilled both cultural and acultural theories.²³ It is not that disengagement and its spin-offs are wholly bad and inner exploration in attunement with nature wholly good (as some 'Romantic' critics of modernity claim), nor vice versa (as some 'Enlightenment' partisans of modernity claim). Modernity's need for self-reassurance arises from the inescapability of *both* horizons. Likewise, the problem of self-reassurance must be dealt with in a way which acknowledges the 'all-pervasiveness' of the modern identity. The problem would be formulated less misleadingly if it were recognized that though a particular hypergood cannot be exclusively pursued without sacrificing others, this does not invalidate that pursued good. In the case of disengagement and attunement, sacrifice is indeed inevitable, but this in itself does not invalidate the good of either.

So far I have been discussing the need for self-reassurance arising from rival conceptions of human dignity, but the point just made about the disastrous consequences of hypergoods also applies to the dimension of moral life which focuses on rights, duties, obligations, and benevolence. Besides its inward orientation to the good, Taylor claims that "the affirmation of everyday life" is central to the modern identity. By this is meant that ordinary life - the life of production and the family - is *affirmed* by moderns as of intrinsic and often overriding importance. The realm of ordinary life is not merely the necessary backdrop to what is really important, for instance being a virtuous citizen or the contemplation of God. It is affirmed or *strongly valued* for its own sake. For Taylor, this hypergood corresponds to the importance moderns place on the avoidance of suffering, and the recognition we give

to persons as *worthy* of benevolence, as bearers of inalienable rights, and as bound in their everyday actions by duties and obligations. But it is also a source of spiritual constraint, particularly in the realm of self-exploration. The reasons for this are no doubt complex, but suffice it to say that not only can allegiance to such everyday life-goods engender suspicion about those who pursue the life of personal self-creation (or, for that matter, the contemplation of God) rather than the life of production and the family, it can also breed subtle forms of persecution of them (motivated by what Nietzsche described as *ressentiment*). On the other hand, exclusive dedication to inner exploration, the word of God, or the pursuit of perfection in either or any other form, can lead to catastrophic consequences in terms of 'life-goods'. This is how Taylor articulates the conflict between art and 'morality', now read as a tension between two poles of the modern identity - the ethic of affirmative self-expression and of selfless benevolence. Both, Taylor suggests, can lead to disaster if uncompromisingly pursued. This much is unquestionable. But it does not follow, he continues, that this undermines the goals of perfection. In my final chapter, I will give reasons for thinking that this latter proposition is indeed questionable, and for reasons which undermine the value of a cultural theory of modernity as espoused by Taylor in general. To anticipate; in order to avoid the errors of procedural accounts of practical reason, Taylor focusses on the identity of a person as a *product*. It therefore gives undue account of the possibility of *disturbances* in the self-formative *process*, and therefore of the motivational constraints on what is worth pursuing.

The affirmation of everyday life is a source of tension within the modern identity for another reason. Its very ubiquity makes it seem that ordinary life is simply a *given*, rather than an affirmation or bearer of strong evaluation. In this sense it encourages the

disengaged naturalist view that the whole notion of strong evaluation is a delusion. And insofar as the modern identity is informed by the axioms of disengagement, it suffers from another source of inner tension; its tendency to frustrate its own articulation *as* an identity. According to the naturalist view, which Taylor convincingly argues exerts a powerful hold on the modern mind, there are no such things as constitutive goods, no 'framework' or 'background' picture which refers to them. While Taylor clearly wants to reject this view - his whole 'cultural' approach to modernity rests on its falsehood - he observes that the *experience* of being without a framework is a common modern phenomenon which itself generates the need for self-reassurance. This it does in the third dimension of the moral life; the problem of the meaning of life as a whole.

It is a peculiar predicament of the modern identity, and hence a distinctive source of its need for self-reassurance, that the *question* of the meaning of an individual life is a problem at all. Moderns need to *look for* an answer to this question, it is the object of a search (or 'quest') which might well *fail*. The problem is not just that the modern might not look hard enough to find this meaning, but that there just isn't a believable one there to be found. By contrast, the problem of leading a full and meaningful life in pre-modern cultures is that of *living up* to the standards defined by the unchallengeable framework. The fear for the pre-modern is failing to do this. But for the modern, the fear is of the meaninglessness of being without a framework at all. This is a danger which is always potentially there, it might happen at any time. When it does, the world loses its "spiritual contour", "nothing is worth doing", there is a "terrifying emptiness, a kind of vertigo, or even a fracturing of our world and body-space".²⁴ The modern is dispositionally vulnerable to this kind of identity crisis, which Taylor suggests is testified by the documented shift in

psychopathologies towards increasing narcissistic personality disorders centering about "ego-loss".²⁵

But if to be without a framework has this experiential consequence, then the claim that we cannot but lead a human life within one, in some sense of 'cannot but', might seem to be vindicated. And this, indeed, is how Taylor argues. To argue this successfully will bolster Taylor's favoured 'cultural' theory of modernity, since if it can be shown that human agents *must* have conceptions of the good life and have resort to a framework of strong evaluations, any theory which denies or fails to accommodate this point (eg. exclusively acultural ones) will at least to this degree be incoherent. I now want to turn to the argument by which he claims to establish that "doing without frameworks is utterly impossible for us...that the horizons within which we live our lives and which make sense of them have to include these strong qualitative discriminations...that living within such strongly qualified horizons is constitutive of human agency".²⁶ If Taylor can successfully establish these claims, he will have gone a long way towards showing that the problem of modernity's self-reassurance *must* be reformulated along the lines we have been considering.

(6.3) The Transcendental Limit of Human Agency: Identity Crises

What needs to be noticed about the above claims is that it is *doing* without frameworks which is alleged to be impossible - that it is *living* within horizons of strong evaluation which makes for human agency. This is an important point to bear in mind, since it tells us something about the nature of the argument which is supposed to lead to these conclusions. According to Martin Löw-Beer's interpretation,²⁷ Taylor cannot coherently be thought to be claiming that the inescapability of frameworks of strong evaluation is either

a conceptual or an empirical truth. It is not a conceptual truth because we can easily imagine a person who prescind from making strong evaluations, or who just "does not give a damn about whether some life-forms are higher and others lower".²⁸ Neither is it an empirical truth, because as we just noticed, Taylor recognizes that there are in fact people who lack an orientation to the good. The necessity of strong evaluations is neither logical nor empirical; the impossibility of doing without frameworks is neither an analytic nor a synthetic truth. Rather, Löw-Beer's interpretation runs, it must be considered as an *existential* impossibility; it would be unbearable *to live* outside a horizon of strong qualitative distinctions. And to show this, Taylor appeals to the phenomenon of an identity crisis - to how the lack of an orientation is lived. While this is indeed part of Taylor's argument, I think he wants to make a stronger claim than is admitted in Löw-Beer's reconstruction of it.

Let us first remind ourselves about what Taylor means by an identity. As a person is a being for whom things matter, so a particular person's identity is what particularly matters for that person - and in both senses of 'particularly'. In the first sense, I am specifically this particular person rather than that, according to Taylor's view, because I take this kind of life to be fulfilling and that kind of life to be empty, or because I interpret this course of action as right and that action wrong, or because I find this species of motivation admirable but that species contemptible. In the second sense, what I find fulfilling or empty, right or wrong, admirable or contemptible, is no small matter, but is of particular or *fundamental* significance to me as a person. My particular identity enables me to answer the question "who am I?". And in answering it, I take "a stand".²⁹ I can always be asked about what really matters to me, what motivates my actions, what gives them their point,

and I can answer by taking a stand in a space which the framework makes available to me. These are the frameworks within which one makes sense of questions about what is good, worthwhile, of value.

To experience radical uncertainty about 'where one stands', to lose this sense of orientation within a horizon of strong evaluation, is how Taylor interprets the phenomenon of an 'identity crisis'. To suffer an identity crisis is to be incapable of telling why a life should be led one way rather than another; it is to be without any sense of discernment between the more or less worthwhile; it is to lose contact with the frameworks within which some life possibilities appear take on more significance than others. The meaning of such possibilities becomes "unfixed, labile, or undetermined".³⁰ As a result, in my identity crisis I feel a kind of vertigo before the question "who am I?". I have access to no resource by which to answer it. But this does not make the question go away. The force of the identity crisis is just that the question of what is really of more or less importance, worthwhile, or fulfilling *demands* an answer even if we are not in a position to give one. Taylor infers from this that frameworks furnish "answers to questions which inescapably pre-exist for us", namely the questions about our identity or the strong evaluations which define where we stand. Only upon the supposition that the question of our identity arises independently of our ability to answer it, does the possibility of *failing* to answer it make sense. Moreover, this failure is experienced both as a *lack* and as an *unbearable* lack. Taylor describes the experience as one of an "acute form of disorientation", "a terrifying emptiness", which presupposes the *absence* of a stand from which to take one's orientation, and implicitly a 'fullness' of being or personhood which depends on it.³¹

It follows from this, Taylor thinks, that frameworks are not something we invent on the basis of matter of fact desires or preferences. This conclusion is opposed to the naturalist view that strong evaluations are dispensable for human beings, that frameworks of evaluative contrasts are an optional extra for human agency. But it is drawn from two quite different arguments which Taylor runs together. First, it follows because the kind of life of which naturalism offers an account is not the kind of life which a human either can lead at all, or can lead without self-mutilation. Naturalism is claimed to be refuted because it "defines as normal or possible a human life which we would find incomprehensible and pathological".³² Taylor seems to be claiming that a life of permanent identity crisis is incomprehensible because we could not recognize it as a fully human life, and that it is pathological because it would be an unbearable life to lead. It thus fails to make sense of a human life as we normally live it; the identity crisis is not comprehensible as a *life*, only as a pathology. This, I take it, is the argument Löw-Beer reads. But there is a second argument, according to which it follows because the phenomenon of an identity crisis is unintelligible on naturalism's own terms. The incoherence of the notion of an identity defined by merely de facto preferences follows because the phenomenon of an identity crisis presupposes a framework of strong evaluation which is *normally* available. As Taylor puts it; "the condition of there being such a thing as an identity crisis is precisely that our identities define the space of qualitative distinctions within which we live and choose",³³ such that these distinctions themselves are *not* something we can choose on the basis of de facto preferences.

Taylor's phenomenological account of identity crises aims to show that the presence of a horizon of strong evaluation is a transcendental condition of human agency. It is a

transcendental condition because without it the leading of a human life would be unintelligible. But our lives are (to a certain extent) intelligible, we do manage to make some sense of them. Indeed we must, as self-interpreting animals, be capable of this. The transcendental condition is what is presupposed in the making sense of life as it is lived. A person in an identity crisis is in crisis as a *living* person. He or she is a person at the limit of personhood.

Is Taylor's account convincing? I see two issues, corresponding to the two arguments I distinguished, which raise doubts. First, one might object that the phenomenon of an identity crisis is not necessarily pathological. Second, one might question whether it must be the case that the phenomenon which Taylor describes is only intelligible under the presupposition of the lack of a horizon of strong evaluations, as Taylor has defined this concept.

Is the phenomenon of an identity crisis - as it is described by Taylor - necessarily pathological? The main problem here is that Taylor, like Habermas, conflates the distinct phenomena of crisis and pathology. That even Taylor relies on this distinction is clear from the contrast between the person who fails to find meaning despite her quest, and the person who fails because she doesn't notice she needs one. A crisis can *collapse* into a pathology, for unlike the latter, the crisis is a state of *transition*. The prospect of a permanent identity crisis would then be equivalent to a life of permanent transition and self-undermining. Such a life may be extremely difficult, or even ultimately unsustainable, but not for this reason pathological. That Taylor fails to give due appreciation to this possibility reflects a deep ambiguity in his phenomenology of an identity crisis; is it human agency *outside* the limits

of personhood, or is it personhood *at its limit*? Only on the former interpretation is it the account of a pathology.

A further objection might be put; that a person might find himself bereft of a horizon of qualitative contrast, but not *experience* this as something "terrifying". Taylor asserts that such a person would not just be 'shallow', 'trivial', merely 'conventional' or the like, rather he would be, as Taylor puts it; "outside our space of interlocution; he wouldn't have a stand in the space where the rest of us are. We would see this as pathological".³⁴ Clearly, *this* sense of 'pathological' is different to the sense in which it is attributable to a person who *affirms* his or her suffering in the grips of a crisis. Indeed, some might claim that far from being pathological, the person in the grips of an identity crisis but who nevertheless 'yea-says' it is *itself* the great challenge of what it is to live a worthwhile life; the requirement would be to tread the tightrope of the crisis without falling into a pathological state. The proper standard of health, someone with this view might propose, is precisely the abandonment of any categories of moral order which obstruct the spontaneous flow of *de facto* desire. According to this view, such natural forces subvert the very purpose of 'taking a stand'. If correct, then the prospects for 'happiness' or the good life would be diminished by thinking of the self as essentially spaced in a horizon of strong evaluation.

This view corresponds to what in (2.4) I called the Nietzschean strategy for 'reconciling' cognition and identity by rejecting the normativity of order. Taylor recognizes Nietzsche's "counter-ideal of the superman and the hypergood of unreserved yea-saying",³⁵ while chastising neo-Nietzscheans who effectively espouse a similar ideal but underhandedly; by denying its status "an ideal, and therefore as a strong evaluation, Taylor claims, they get

caught up in a performative contradiction. But the 'ideal' espoused here is undertaken through the exploration and affirmation of precisely limit experiences, those which push back the bounds of reflective - including hermeneutic - subjectivity. Since these are experiences of *self-dissolution*, the neo-Nietzschean would retort, they hardly look like good candidates for an 'identity' in Taylor's sense.³⁶ The neo-Nietzschean might suggest further that Taylor's implicit dismissal of the 'identityless' self as pathological rests unwittingly upon a reassertion of the value placed by moderns on the affirmation of *ordinary life*. In (8.2), I will indeed argue that Taylor wants to have his cake and eat it; that his argument for identity as a stand in the space of strong evaluations is incompatible with his interpretation of how the 'ideal' of 'non-identity' is *articulated* in modernist, 'epiphanic' art.

Concerning the second issue, one might want to object against Taylor that our account of an identity crisis need not appeal to strong evaluations. It helps Taylor's case against the first objection that a crisis, as we saw in (1.2), is something which is *suffered*. But then what needs accounting for, this next objection runs, is the causal process which explains it as a physical phenomenon. Such a physicalist explanation would trace back the cause of why I suffer when I am asked the particular, weighty, existential question, "who am I?". But it would also seek to explain why someone else does *not* suffer when asked this question, and also why that person might suffer, even to the point of crisis, when asked apparently much more trivial questions. In the next section, I will develop this objection by considering a position which combines this naturalist bent of thought with the advocacy of the kind of self which "yea-says" the undermining of its own identity.

(6.4) Pragmatist objections

Taylor claims that for a person to live outside a framework of strong evaluations is unintelligible. To be a person is always to be a more or less of a person, it is to live a life which can be recognized by the person as satisfying or failing by the normative standards of the framework which must be supposed to hold independently of the person's de facto desires. This view is proposed in opposition to 'naturalism', which can be defined as the doctrine that, rather than being a transcendental condition of human agency, frameworks of strong evaluation are an optional extra for human beings. Although Taylor explicitly attacks utilitarian and 'emotivist' conceptions of human agency, I want now to consider as an objection to this account an alternative - and I think more challenging-version of naturalism; the pragmatism recently espoused by Richard Rorty.³⁷ Rorty's position is more challenging because while it accepts the hermeneutic claim that persons exist within frameworks of self-interpretation, these self-interpretations are construed not, *pace* Taylor, transcendently and ontologically, but naturalistically and pragmatically. Under this construal, what Taylor argues is outside the limits of the human, Rorty offers as an ideal character type of a 'postmodern' culture.

As a naturalist, Rorty maintains that there is nothing more to say about the self once its de facto desires, beliefs, hopes etc., have been fully accounted for.³⁸ There are, however, constraints on what particular desires and beliefs can be ascribed to a *person*. Following a cue from Davidson, Rorty holds that to ascribe a particular desire or belief to a person is always to ascribe it as part of an internally coherent and prima facie plausible set or web of desires and beliefs. The particular desires and beliefs which make up a person are intelligible in virtue of fitting with others. But often there appears to be a lack of fit. For

a naturalist, this would be the case where the desirability-characterizations which Taylor calls strong evaluations stand opposed to one's matter of fact desires. And as I discussed in (3.2), it is also the case between 'unconscious' and 'conscious' desires and beliefs. To render the lack of fit between the unconscious and the conscious self intelligible might therefore point the way to a naturalized self which can dispense with, while account for, strong evaluations. This, in effect, is the path followed by Rorty.

We know that people are often incoherent in their desires and beliefs, and sometimes radically so. How is this phenomenon intelligible given the supposition that a person is an internally coherent cluster or web of desire and belief, the parts of which are only ascribable to a person insofar as they are ascribable to such a coherent whole? Since, Rorty supposes, it is an incoherence of *agency* which is the problem here - that it is a desire or belief of a person which poses the problem of intelligibility - it won't do to reduce the particular incoherence to a person-neutral (eg. neural or endocrinal) mechanism. Rorty's move is to posit *another* self, or what he calls a 'quasi-self' or 'person-analogue', as constituted by another net of desires and belief of which the *prima facie* incoherent desire or belief is reinterpreted as a coherent part. Again as a naturalist, Rorty insists that each of these quasi-selves "is part of a unified *causal* network", though this is not the unity of a person.³⁹ This causal network, "the machine that each of us is", is the product of the very particular and idiosyncratic contingencies of the person's upbringing or natural history. In the course of this history, Rorty argues, we form several internally coherent but mutually incompatible nets of desire and belief, only one of which is available to introspection at any one time. This much, Rorty proposes, we learn from Freud. But according to Rorty's reading, there is a lesson about moral reflection to be drawn from Freud's conclusion - one

which signals a challenge to Taylor's view on the relationship between the self and the good.

As we have seen, Taylor's concept of identity turns on the idea of qualitative contrasts. His thesis is opposed to the naturalist levelling of these contrasts. Now Rorty is also opposed to a certain *kind* of levelling; that which is perpetrated by *reductive* theories of agency. Rorty does not think that the meaning of qualitative contrasts can be reduced to some homogeneous operation like ego-utility maximization, 'qualia' association, evolutionary benefit, or the like. For Rorty, this kind of reductive naturalism is rightly criticized by Taylor. The levelling proposed by Rorty is of a quite different sort; not the reduction of the self to some inarticulate mechanism, but its *decentering*. For Taylor, *de facto* desires, motivations, and actions, await assessment as higher rather than lower, admirable as opposed to base, truly worthwhile against superficial, and one's identity is defined by what counts as in the former categories. The stand which one takes in the light of these characterizations *centres* the self. From this centre, judgment can be passed on the marginal and the superficial. Now Rorty would claim that his view can accommodate generic distinctions of this kind, but rather than hierarchically opposing a true or authentic self to a *de facto* net of desires, he construes the former just as a *different* self (or network of desire and belief). This is not to say that strong evaluations are *mere* preferences, except in the sense in which we can talk of 'mere' persons. Qualitative contrasts or strong evaluations should therefore be taken as contrasts between different selves, rather than between 'the self' and 'de facto' desire. What Taylor calls the self would then be understood not as what stands independently of *de facto* desire, but rather as another pattern or web of desire and belief which stands *alongside* dispersed sets of desires and beliefs.

It only seems otherwise, Rorty might argue further, because at any moment of reflection there is just one self open to introspection. Accordingly, what counts as more or less worthwhile will vary between the different nets of desire and belief (or quasi-selves) which make up any particular person. And this means that the question, inescapable on Taylor's view, of what counts as of *fundamental* importance drops out of focus.

The question only has a focus, Rorty would argue, where the point of moral reflection is to unearth the "essence" of personhood, the "core" self, the "truly" human, the "real" me, etc. And this is how Rorty might view the role Taylor's strong evaluations. As answers to the question "where do I stand?", strong evaluations now look disingenuous. For given the multiplicity of quasi-selves which make up the person, what "This is where I stand" means will turn on which "I" is asking the question. Further, the question is loaded in favour of Taylor's view in virtue of the fact that only a certain kind of "I" *will* ask the question; it will be raised by a reflective self with access to a certain language of self-interpretation. It is also the self which defines the essence of the human *as* self-interpretation, as addressing the interpretative question "who am I?". Against this, Rorty favours the view - which he takes to be encouraged by Freud - that we "see ourselves as centerless, as random assemblages of contingent and idiosyncratic needs rather than as more or less adequate exemplifications of a common human essence", even if that essence is taken to be self-interpretation.⁴⁰

So far I have been reconstructing Rorty's objections to Taylor as a naturalist, but this thought leads us to the alternative model of the self and moral reflection he offers. As a pragmatist, Rorty insists that the *de facto* desires, hopes, and beliefs of a person are never

fully accounted for; they are ever open to reinterpretation and redescription.⁴¹ One can pragmatically adapt to the contingencies of selfhood by self-creation through self-redescription. The point of moral reflection would then take a corresponding turn; it would shift from the grounding of an identity ('where I stand') to the exploration of the different, hidden 'quasi-persons' which make up a self. Rather than searching for one's "true centre", moral reflection would be directed towards becoming "acquainted with these unfamiliar persons".⁴² The question "who am I" need not then be "where do I stand?", but "who causes me to have my strong evaluations?". The task would be to reconcile in conversation what splits us up 'into incompatible sets of beliefs and desires'. If we think of strong evaluations as analogous to the voice of conscience, discriminations between the significant and trivial will be taken as part of "*just* another story", like Freud shows us that the story of the superego is on a par with the stories told by the ego and the id. In abandoning the idea of a single story holding them all together, so the idea of an overriding narrative of a life *quest* loses its hold. The idea of a unifying, self-perfecting quest should be dropped in favour of *ad hoc* narratives tailored to the contingencies of individual lives. Such small narratives are "more plausible because they will cover *all* the actions one performs in one's life, even the silly, cruel, and self-destructive actions"; they will be without "heroes or heroines".⁴³

To refrain from asking Taylor's question of identity requires that a different attitude be adopted to the self, and conversely that the self is better thought as something for which a different attitude is apt. For the pragmatist, maturity involves substituting recognition of chance as "not unworthy of determining our fate" for the aspiration to see things "steadily" and "whole".⁴⁴ From a conception of ourselves as "tissues of contingencies" can be derived

an imperative to self-creation. According to Rorty, such a pragmatic attitude would encourage us to become more "ironic, playful, free, and inventive in our choice of self-descriptions". Presumably, identity crises would then be open to re-description. If there were to be no longing for a centre, if I were to take an ironic attitude toward myself before the question "who am I?", the anxiety of being without an orientation would be defused. Indeed, it *ought* to be defused, not only because the idea of a centred self is deluded, but because it is a hang up from a culture which has lost its use. In Rorty's postmodern culture, the very phenomenon of an identity crisis would be unintelligible. It would be viewed as a quaint relic from the foundationalist past. The self would have lost its *tragic* dimension.

Taylor might reply to Rorty's alternative conception of the self by returning the following related questions; (1) how pertinent is Rorty's representation of the challenge of Freud? (2) Does Rorty's alternative conception of the self really do without a framework of qualitative contrasts? (3) If it does, is it a coherent conception of a livable life? The issues raised by these questions overlap, but it will help to deal with them separately.

(1) We have seen how Freud's splitting of the self might be seen to change the force of the distinction between the admirable and the contemptible, the higher and the base. How seriously does this impinge on Taylor's position? On the one hand, Taylor is very much aware that the disengaged, controlling, disciplining, instrumental stance of the self, which according to Rorty is represented by the Freudian ego, is criticizable by appeal to more spontaneous, playful action orientations. Insofar as 'unconscious' impulse can be thought of as the source of such orientations, then it is not the case that Taylor identifies the

generically 'lower' with 'unconscious' motives, and his position seems unthreatened by Rorty's Freudian model of the self. On the other hand, however, Taylor's *justification* for the inescapability of horizons of strong evaluation turns on the *conscious* state of the person in an identity crisis; strongly evaluated motives must be conscious, as must be the 'stand' by which, for Taylor, the self wins back its identity. So it does seem as if there is something question-begging in Taylor's argument, since self-interpretative question 'who am I?' *presupposes* a reflective and consciously situated 'I' for whom the question is inescapable. But while this objection holds at the transcendental level of justification, it by no means follows that on the basis of a naturalistic and pragmatic construal of strong evaluations, Rorty's Freud offers us a superior model of the self or of moral reflection than Taylor does. In the remainder of this section, I shall go some way to justifying this claim.

Take Rorty's insistence that Freud enables us to relativize discriminations between the significant and the trivial - of the kind which strong evaluations perform - to stories constructed in an ad hoc, pragmatic manner, to suit the particular contingencies of a person's self-development. For Rorty, such 'small narratives' have the advantage of covering *all* one's actions, even those which are silly and trivial. This view gains its appeal from its opposition to practices of 'moralizing' which are insensitive to the particular life-histories of agents; to what, as we saw in (3.1), Williams calls the institution of blame. But there are several problems with Rorty's presentation of the source of that opposition. First, the psychoanalytical narrative *uncovers* the significance behind what is only *apparently* trivial. But this significance-disclosing *transformation* of the meaning of actions requires that it is *already accepted* that what the general psychoanalytic theory covers is of fundamental concern; namely, the roots of psycho-sexual development in early childhood

experience. Clearly, insofar as their potential for articulating discriminations regarding the worthwhile and the worthless in a life-history goes, stories in this domain are not on a par with all others. The new language of self-interpretation may enlarge the scope for significance, but it does not undermine the distinction between the significant and the trivial. Second, the morally charged distinction between the significant and the trivial is presupposed in the conditions of acceptability of the 'small narrative'. One only *begins* the 'conversation' with other quasi-selves because *these* particular feelings do assume a highly significant role in the leading of one's life. As we saw in Habermas's interpretation of Freud, these are grounded in a 'passion for critique'. But this passion for *critique* is motivated by the need to change for the better, a motivation which would be unintelligible in Rorty's scenario, and which is captured in Taylor's concept of strong evaluation. Third, Rorty makes it seem as if the shift to new vocabularies of self-enlargement is as contingent as the life-histories which they tell. But this either leaves the question of why one particular narrative is preferable to another unanswerable, or if pragmatic criteria are appealed to, it leaves the question unanswered, since the question of what counts as useful - Rorty would agree - is internal to the vocabulary of self-interpretation, and is thus begged by appeal to pragmatic considerations.

(2) The crucial point for Taylor - and the decisive advantage of having the conceptual resource of strong evaluations - is that one can acquire 'self-enlargement' not only through exploration and conversation with one's 'quasi-selves', but also through *critical* conversation with them. They are criticizable, he would maintain, because their contrasting orientations are strongly evaluated. For itself to be *affirmed* as good, 'self-enlargement' must exercise its pull in contrast to something less humanly fulfilling; for example, the (typically

parodied) life of the self-responsible, duty-bound 'Kantian man'. In contrasting the two types of moral reflection and character type, Taylor can reply, Rorty is exploiting a qualitative contrast.

Indeed, there are points at which Rorty comes very close to Taylor's view. We have seen that Rorty shares with Taylor an acceptance of the importance of self-interpretations to the extent that Rorty asserts that a self (a net of desires, hopes and beliefs) becomes "richer and fuller" by developing richer and fuller ways of formulating one's hopes and desires. The main technique of self enlargement and development is the "acquisition of new vocabularies of moral reflection".⁴⁵ By this phrase, Rorty means "a set of terms in which one compares oneself to other human beings", for example as "magnanimous", "decent", "cowardly", "epicene". These terms feature as answers to questions of identity like "What sort of person would I be if I did this?". Rorty asserts further that the availability of a richer language of moral deliberation makes us "more sensitive and sophisticated than our ancestors or than our younger selves".⁴⁶ But to acknowledge this relative deficiency in 'sensitivity', to answer the question of identity as "I would be more sensitive if I did this", Taylor could claim, is to cede the point about frameworks of qualitative contrast. And the point is generalizable to any attribute through which one compares oneself to other human beings, real or imaginary.

These considerations suggest an incoherence in Rorty's naturalistic reinterpretation of strong evaluations. His holding operation is to propose that an ironic attitude be adopted to any supposed self-improvement which is conditioned by the availability of a new language for self-interpretation. The justification for Rorty's 'ironism' lies in the thesis that the claim

that one is *really* more (or less) sensitive, decent, maganimous, etc. than one's younger self, presupposes commitment to an 'essence' of the self - an absolutely non-contingent ideal of personhood - which one can fail or succeed to realize. But this desperate reduction of alternative positions to 'Raving Platonism'⁴⁷ betrays the weakness of Rorty's own. Taylor's construal of qualitative contrasts does not rest on such a notion of an 'ideal' self. All that it requires are terms of comparison between two concrete selves or actions.

(3) If it is the case that the contingency of selfhood, as Rorty describes it, is inconsistent with the applicability of the concept of strong evaluations, is it a self which is livable? The question will seem bizarre to someone who is convinced by Rorty's view, since his claim is not only that such a life is livable, but much more strongly, that one *ought* to lead it, at least insofar as one aspires to the values of a liberal culture. Evidence from psychoanalytical studies suggests, however, that it is precisely the *scarcity* in resources of identity-carrying horizons of significance which, at least in part, makes the leading of life unbearable in actually existing 'liberal' culture.⁴⁸ Taylor thus has the option of retorting that *as far as we know*, an individual needs the stability of an identity *even to pursue the project of self-creation*. This kind of argument does not have the 'transcendental' force of Taylor's justification for the inescapability of strong evaluations, but it does put the onus of argument back on those who deny his claim.

Furthermore, and this is a point to be learned from Habermas's interpretation of Freud, the project of self-determination presupposes a condition of mutual recognition. What is massively conspicuous in its absence throughout Rorty's Freud is the phenomenon of *resistance* and its link with the therapeutic goal of psychoanalysis.⁴⁹ But resistance can be

explained, we saw, as an operation which makes bearable an otherwise unbearably traumatic mental event which results from the distortion of conditions of mutual recognition. There is no space in Rorty's account for the role played by *power* in the *disturbance* of contingent self-formative processes. I will suggest in (7.3) below, that it is a space which Taylor must also make available.

There are other questions which, as yet, remain unanswered in Taylor's defence of Rorty's pragmatist objections. Taylor insists on the importance of dropping an identity when the moment is appropriate, but how are we to determine this moment? Or better; how are we to rationally justify the *transition* from one self to another? The role of Taylor's phenomenology of crisis is to *establish* the inescapability of frameworks. In the next chapter, I shall consider how the phenomenon of an identity crisis is used to account for rational *change* between frameworks. How, in other words, can the 'turn towards other models' on which I quoted Taylor at the beginning of (6.2) be rational? But before that, I want to draw together some parallels between the role played by a concept of crisis in the proposals for critical reflection put by Habermas and Taylor.

(6.5) Crisis and Critical Reflection

I have been examining some of the conceptual resources which Taylor wants to make available to a cultural theory of modernity. These are offered as an alternative to those which equip acultural theories, of which Habermas's is targeted as one. I will leave an overall assessment of their the relative merits of their competing conceptions of modernity, and especially of their articulation of the problem of self-reassurance generated by the tension between cognition and identity, to my final chapter (chapter eight). But from our

discussion so far, certain thematic convergences come into focus, to which it is now timely to draw attention.

The first theme of convergence concerns both Habermas's and Taylor's application of a *reductio ad absurdum* form of argument to the kind of 'norm-free' human conduct which is the target of the critique of instrumental reason. Instrumental reason concerns the efficiency of means to non-normatively evaluable ends, and the critique of instrumental reason has to demonstrate that there are good reasons for acting upon norms to which instrumentally rational actions can be brought to account. For Habermas, this demand is equivalent to that of showing that there are good reasons for acting 'communicatively' rather than 'strategically'. For the strategic actor, other human beings are of significance solely to the extent to which they feature as manipulable means to ends which the strategic actor has himself decided; they are of value just to the extent to which they are instrumental towards the satisfaction of his desires. For Taylor, the demand is equivalent to establishing that the matter of fact, naturally 'given' desires upon what Taylor calls the 'weak evaluator' acts, are accountable to independent standards - what Taylor calls frameworks of 'strong evaluation' - of what is desirable for human beings. Both Habermas and Taylor argue that there is something 'inescapable' about the norms by appeal to which a critique of instrumental society can be undertaken; in the former case, these are norms which are presupposed in communicative action, in the latter, they are norms which are presupposed in frameworks of strong evaluation. But neither the indispensability of communicative action, nor the inescapability of horizons of strong evaluation, are strictly speaking logical requirements. Rather, they are claimed to be *the practical demands of a worthwhile existence*. Both the life of the strategic actor and the life of the weak evaluator are claimed

to have consequences the recognition of which, if anticipated by those who advocate them, would render their advocacy absurd. The absurdity here lies not in a logical contradiction, but in the espousal of a form of life which generates unsustainable identity-crises; the practical undermining of the agency which would choose to lead such a life.

Since for an agent to suffer an identity-crisis is for the conditions of his agency to be undermined, both Habermas and Taylor can appeal to a sense in which strategic action and weak evaluation fail to satisfy the transcendental conditions of human agency. Again, such conditions are 'transcendental' in a practical sense, insofar as failure to satisfy them is inconsistent with a normatively secure, worthwhile human existence. Habermas's strategy is to trace these conditions back to the context of 'intersubjectivity' only within which human identities can emerge in a non-distorted form. This context he calls the 'lifeworld'. Accordingly, the critique of instrumental reason can take its departure from the norms which are presupposed in the undistorted reproduction of the lifeworld, where the carrier of this reproduction process is communicative action. His claim is that unless the self-formative process of individuals and collectives is mediated communicatively rather than strategically, socially induced pathologies of identity which follow a systematic pattern ensue. Habermas's central claim about communicative action, then, is that the strategic or 'norm-free' mediation of subjects is incompatible with the self-formation of human beings in a non-pathological form. For Taylor, on the other hand, the non-availability of a framework of qualitative contrasts within which to characterize what is desirable has a similar, though even more radically devastating effect on the human subject, since what it is to live a worthwhile human existence is defined within the framework. Habermas's lifeworld and Taylor's frameworks share this fundamental feature; they are the horizon in

virtue of which human agency has an orientation, (and therefore the condition for the possibility of human agency at all).

A second, related theme of convergence, lies in the appeal made by both Habermas and Taylor to the phenomenon of an identity-crisis as a *reductio* against a certain kind of value scepticism. Habermas's value sceptic holds the view that rationality is exclusively instrumental, that there is no good reason for an agent to be constrained in his ego-utility maximization by supposedly intersubjectively valid norms of action. For an agent to recognize such restraints, Habermas holds, is for him to acknowledge the validity of the moral point of view, where the moral point of view represents the perspective of the impartial participant in practical discourse. This point of view has validity, according to Habermas, in virtue of the pressure towards universalization which is built into the procedure for reaching a rational agreement. The moral scepticism targeted by Habermas is, in the first instance, the view that this procedure is insufficient for generating valid and universalizable principles of justice. Habermas replies to this that as a rationally justifiable claim, it is subject to those very universalizable norms which it simultaneously presumes to reject. But to the sceptic who does not claim such warrant - whose assertion is a strategic rather than a communicative act - Habermas rejoins by pointing to the unsustainability of a form of life in which *that* were universalized. In the first stage of the anti-sceptical argument, the unavoidability of the normatively charged presuppositions of communicative action is 'law-like', in the second - and I have argued more fundamental - stage, the unavoidability is 'fate-like'. The 'disenchantment' of this fate turns partly on the internalization of a theoretical interpretive perspective on self-formative processes which provides the agent with non-instrumental reason (and motivation) for acting. Taylor's moral

sceptic, on the other hand, is the 'naturalist'. The naturalist also holds that rationality is exclusively instrumental, that there is no good reason to be constrained in the satisfaction of my de facto desires by considerations of the intrinsic worth of those desires. Both Habermas and Taylor claim to reduce the respective moral sceptic to the absurdity of espousing a course of life the consequences of which are crises of identity.

Third, moral scepticism of both the above kinds, Habermas and Taylor agree, is the product of a more general self-misunderstanding of modernity. For Habermas, the source of this misunderstanding is the 'philosophy of consciousness' and its associated paradigm of 'subject-centred reason'. For Taylor, it is a false philosophical anthropology of 'disengagement'. Both theorists trace the respective confusions back to the Descartes/Kant canon of modern philosophy. The main confusion of the philosophy of consciousness is that it construes rational thought and action 'monologically', on the basis of the reflecting subject's transparent relation to itself. This generates the self-misunderstanding that the growth in 'rationality' which characterizes modernity is itself monological or 'instrumental' in form, breeding the sceptical conclusion that in a world of subjects relating to objects, there can be no good reason for acting morally. The central confusion of the anthropology of disengagement, on the other hand, is the view that the only truths are those available to the observer who, by adopting the kind of scientific methodological position *to* the world, neutralizes his engaged stance *within* the world. This makes it seem as if the only genuine form of reasoning is one which follows the rigorous procedures of modern natural science, and that in the modern world genuinely 'neutral' values have been discovered or are aspired to. But it also generates the sceptical outlook that concerning 'values', nothing can be found which affords of the same kind of certainty and interpretive unanimity which

characterize 'facts'.

Fourth, the abstract errors of modern philosophical scepticism reflect on the concrete need for modernity's self-reassurance. For both Habermas and Taylor, as I am interpreting them, this need arises from a tension between cognition and identity. Habermas construes this tension as resulting from a one-sidedly rationalized lifeworld along its 'cognitive/instrumental' dimension. The lifeworld is what provides the resources for identity-formation, it is the locus of the claim to normativity which give agents reasons for acting non-instrumentally. Individual and collective agents emerge from a lifeworld which precedes them, and they reproduce the lifeworld through their actions. However, under the pressure to maximize the instrumental efficiency of actions coordinated for the purpose of reproducing the material resources of modern society, the capacity for controlling nature and social systems within the lifeworld becomes excessively exploited. Since science and technology can be harnessed to increase the efficiency of an ever-growing system of material reproduction, the 'value-sphere' in which cognitive/instrumental rationality is condensed undergoes a corresponding expansion. But this happens at the expense of the value-spheres of morality and art, since the forms of rationality appropriate to them obstruct, rather than promote, the expansion of the capitalist economic system. The integration of actions in the latter is 'norm-free', and hence in tension with human identity. Taylor agrees that value-scepticism is indeed a symptom of modernity understood as a merely 'instrumental society', where the tendency to become an instrumental society is central to what generates the need for self-reassurance. But for Taylor, the tension between cognition and identity arises from a misguided generalization and over-emphasis on the goods which accrue from taking a disengaged stance to the world, two of the most

important being scientific knowledge of nature and its technological control. So although for Taylor the tension between cognition and identity can be expressed in terms of the dominance of the cultural value spheres which correspond to Habermas's cognitive/instrumental dimension of the lifeworld, these are theorized as normatively significant themselves, rather than as norm-free.

The fifth thematic convergence I want to highlight, is the affirmation on the part of both Habermas and Taylor that recognition of the linguistic constitution of identity, and the unseverable embodiment of rationality in language, is the first step along the road to thinking clearly about how cognition and identity might be reconciled. Habermas and Taylor share the view that self-understanding can only be achieved in the medium of language, that language must be understood as a rule-governed activity, and that the rules which govern linguistic meanings are public. They concur on the view that language cannot be given a foundation by appeal to something to which it 'corresponds'; be it the self-transparency of 'sense-certainty', or language-transcendent metaphysical 'facts'. They agree that just as language is the medium in which identities are forged, and thus that the possession of a language is the sine qua non of human agency, so is language the medium of cognition, and thus the possession of a language is the sine qua non of human knowledge. Both appeal to the fact that the same linguistic vehicle carries both identity and cognition as the point of departure for thinking how the two might be brought to a reconciliation. The motto that "all philosophy is a 'critique of language'"⁵⁰ is apt for both their conceptions of critical reflection.

But after this common starting point, their paths diverge radically. The point of departure,

and what essentially separates their respective 'cultural' and 'acultural' frameworks for critical reflection on modernity, lies in the different roles each attributes to language. For Habermas, language serves a function which all societies must perform; social integration. The stability of any society requires that symbolic resources are made available which secure both the individual and collective identities of its people. In modern societies, religion is unable to serve this function. This is the byproduct of a learning process through which objective scientific knowledge is separated off from moral and aesthetic values. As a learning process, it can be reconstructed as a process of rationalization - or as sequences along a logic of the development of communicative competence. This idea lends itself to the thought that ideal procedural standards of rationality - standards which are objective in virtue of being unavoidable presuppositions of an unavoidable mechanism for co-ordinating actions - can be tapped by the critic. If modernity is to preserve its claim to rationality, it can not regress to previous stages; for instance, to norms which are 'substantive' in being part of the fabric of the world. The task of critical reflection, then, is to reconstruct the norms which are presupposed in language itself; to show that resources are available for the integrity of modernity in virtue of its mediating mechanism. The burden of Habermas's theory, then, is to show that there are rationally binding, non-instrumental norms which are immanent to language itself. But it is also obvious that there are 'strategic' uses of language; forms of linguistic interaction which are instrumentally rational. Therefore Habermas takes it upon himself to establish that there is an 'original' use of language which is 'communicatively' rational. Further, since language is the source of normativity, he can propose that the communicatively distorted mediation of human beings is responsible for the crises which generate the need for self-reassurance.

For Habermas, then, the 'critique of language' takes the form of a theory of communicative competence, which attempts to reconstruct a logic of development, and which provides standards against which systematically distorted communication can be diagnosed as crisis-generating. In contrast to this, the critique of language which Taylor undertakes is focussed on the ideals which language makes available for self-interpretation. Self-interpretation is theorized as the rendering articulate of a moral ontology which is disclosed by language. This difference in philosophical perspective is the source of what appears to be the most intractable difference between Habermas and Taylor; for the latter, Habermas's commitment to a procedural conception of rationality proves that he remains in the yoke of disengaged anthropology, for the former, Taylor's alternative falls foul of the logic of communicative competence, requiring an impossible task of 'unlearning'. In chapter eight, I will assess the applicability of the charge that Habermas's proceduralism betrays commitment to a philosophical anthropology of disengagement, and I will introduce the radical sense in which, for Taylor, under conditions of modernity, ideals for self-interpretation can be disclosed only *through* language in certain *form*. But in the next chapter, I turn to what, for Habermas, are the two major obstacles to Taylor's conception of critical reflection: its assumption of 'metaphysical premises' which are incompatible with the cognitive or 'learning' achievements of modernity; and its tendency towards conservatism in its manner of reflection upon morality and ethical life.

CHAPTER SEVEN: TAYLOR, LANGUAGE AND CRITIQUE

(7.0) Introduction

Even if we accept Taylor's thesis that frameworks of strong evaluation are inescapable for human beings, and that the task of a cultural theory of modernity is to articulate the complex framework which constitutes the modern identity, it remains for us to be shown how the modern identity can thereby be *reassured* about itself. If cultures (and individuals) differ in the strong evaluations they carry, and if one's allegiance to a culture is itself the expression of a strong evaluation, then what reason is there for being reassured that the evaluation through which one identifies oneself is more defensible than others? Won't the reason be relative to the particular framework one inhabits - the particular 'stand' one has adopted? And if this is the case, where does it leave the status of the evaluative claim? Does it make sense to understand it as a 'truth-claim'? Or must it be construed as a claim whose truth is relative to the particular framework in which it finds expression, and hence not strictly speaking a truth claim at all? And if it must be so construed, does this not undermine the viability of the cultural theory of modernity itself?

In this chapter I want to consider how Taylor attempts to disarm the relativist objections which appear to beset his cultural theory. In chapter one, it was noted that in calling for a debate between cultural and acultural theories of modernity, Taylor is aware that the claims of the former kind of theory lend themselves to a relativism from which, if they are to be acceptable, they must be redeemed. Cultural theorists of modernity, Taylor observes, have been unable to dispel the appearance of espousing a self-defeating relativism. Recognizing that his preferred cultural theory cannot embrace this implication, Taylor needs

to establish the sense in which changes in the framework of self-interpretation can be rational without resorting to a culture-neutral criterion of rationality. I consider how Taylor attempts to do this in (7.1). Practical reason, Taylor proposes, can account for how changes in identity can constitute an epistemic gain by taking a narrative form. Narratively structured practical reasoning is substantive in virtue of carrying the content of rich conceptions of human identity, rather than prescind from that content in the manner of the procedural model of practical reason. But it is also substantive in that its conclusions are affirmed as admitting of truth. Taylor thus needs to establish the thesis that concerning strong evaluations, there is a truth to the matter. I reconstruct this stage of Taylor's argument in (7.2), by focussing on the internal relationship claimed between a favoured cultural theory and moral realism. The key to this argument, I suggest, lies in the *explanatory* power assumed by identity-expressive, world-disclosive language. Only a vocabulary enriched by the kind of concepts through which life is lived is sufficient for certain explanatory purposes, Taylor claims, and therefore capable of featuring in the best accounts of the world. By bringing together the concepts of explanation and practical reason within a world-disclosive or identity-expressive model of language, a cultural theory can thereby bring to self-clarification its potential for rationally grounded critique.

But this move is not without its problems. Also in chapter one, I remarked that Taylor's case for a cultural theory of modernity can be taken in a strong or a weak sense. The weak claim is that the change in scientific outlook at the onset of modernity can be described in acultural terms as opening up truths which 'we have come to see'. The strong claim is that *even this* opening up of truths is best described in terms appropriate to a cultural-theoretic framework. This ambiguity, I propose, lies in the equivocal status ascribed by Taylor to

the 'absolute conception' of reality. The sense in which we can talk about truth and the real, Taylor sometimes advocates, needs to be distinguished from the conception of truth and reality available to the disengaged enquirer. Once the cultural theorist does this, Taylor claims, the case can be pressed for moral realism. But at the same time, by drawing upon the rational form which the resolution of *crises* within scientific traditions (or 'paradigms') takes, he advocates an assimilation of the conceptions of truth and reality available to any enquirer insofar as truth and reality are what are disclosed in the most rational account. After indicating how Taylor's views issue in a tension between scientific and moral realism, I turn in (7.3) to some objections to the role played by language within the interpretative paradigm of critical reflection carried by cultural theory. In (7.4), I return to Habermas's concern that attempts to bring the claim to normativity to self-clarification on the basis of an expressivist, world-disclosive theory of language bear a conservative prejudice. This will also allow me to consider how a problem of mediation between morality and ethical life arises within a cultural theory. Following up a distinction made by Sabina Lovibond between 'empirical' and 'transcendental parochialism', I suggest a way of quelling Habermas's worry about the conservative tendencies of the expressivist view.

(7.1) The Narrative Form of Practical Reason

As we have seen, Taylor's preference for a cultural theory of modernity rests on the claim that the conceptual resources available to it - the notions of 'identity', 'strong evaluation', and 'the good'- are indispensable for making sense of human life as it is lived. Human agency, Taylor claims, involves the capacity to take a stand on the question 'who am I?' or 'what really matters to me?'. In other words, it requires having an orientation in moral

space, one which is mapped by a framework of qualitative contrasts. Against this thesis, I raised two kinds of objection. First, that the *question* of identity is only inescapable for a certain kind of self; that of a reflective subject whose self-interpretation is of fundamental concern. Second, I suggested that the 'stand' defining one's identity may be entirely provisional, forever open to further unmasking and change. To counter this objection, Taylor can reply by appealing to another conceptual resource available to his cultural theory, one which I have not yet considered. For in addition to taking a stand in moral space, Taylor asserts, a human being cannot but *move* in moral *time*. This moral or human time is that which is articulated in narratives. Taylor then exploits the rationality of narratively articulated changes in identity for dealing with the problem of the *relative* merit of modern and other identities. By appeal to a phenomenology of those changes in a life which constitute moral *growth*, Taylor extracts a model for comprehending what makes the rational critique of modern and other cultures possible. But if Taylor is able to do this convincingly, then he will have effectively disarmed Habermas's worries about the rationally toothless critical powers of a cultural theory of modernity. I will return to this point in a moment. But first, a few words on why it is thought that a human life *must* be understood narratively.

Taylor thinks this follows from the necessity of our orientation towards the good. The inescapability of the question 'who am I?' (or 'where do I stand?'), is also the inescapability of a question which can be raised at *different* times. It is thus bound to the further question of 'who am I *becoming*?' (or 'where am I *going*?'). At any moment, the orientation which defines my identity is something at which I have arrived, and grasping the significance of it requires an understanding of how I got there. But this is the kind of

understanding, Taylor insists, which can only be grasped by placing one's life in a story. Such a life story will always be open to revision, and its revisability in the light of future changes and becomings alters the meaning of how *they* were arrived at. To understand my life in a story, then, requires that it is a story of a *whole* life. This is also something which my orientation to the good requires; "We want our lives to have meaning, or weight, or substance, or to grow towards some fulness...But this means our *whole* lives".¹ And a whole life is only comprehensible in narrative form.

In virtue of its temporal constitution then, human agency is conditioned not only by an orientation towards the good, but also by the possibility of moral growth. We have already seen that for Taylor, one's fundamental orientation to the good is determinable by reference to the particular hypergood to which one gives one's allegiance. The question then arises of how the hypergoods which define an identity are to be deliberated over. This is the question which, according to Taylor, a theory of practical reason must answer. It must make sense of the possibility of a rational *transition* from one identity to another; it must give an account of the kind of move which constitutes moral growth. In chapter five, we saw that Habermas too wants to open up a space for rational deliberation concerning questions of the good life, though only in the domain of what he calls therapeutic discourse. However, it was noted that clinical intuitions, in not admitting of a universal core which is unconditioned by culturally specific horizons of self-interpretation, cannot attain the strict claim to rationality of judgements of right or justice. For Habermas, the rationality of clinical intuitions is compromised by the absence of a universalizable criterion of the kind which the idealized symmetry and reciprocity conditions of communicative action afford. But by posing the question of rational practical deliberation in terms of the *transition*

between action-guiding orientations, the absence of such universalizable criteria need not look so compromising. Furthermore, by posing the problem of practical reasoning in this way, the motivational deficit of agents whose obligations derive from the idealizations of the moral point of view is covered, without thereby forfeiting the rational force of the practical insight. Of course, the question of what is right or just to do remains a subject of practical reasoning, insofar as actions (or norms) are an expression of a particular identity which it is better to be (or which merits allegiance), on account of its respect for the obligation to satisfy universalizable needs. But obligations are only *worth* respecting insofar as they orient the agent to the good. The correctness (whether it counts as 'growth' or not) of the particular practical deliberation is determined by the comparative superiority of the positions on either side of the move, rather than by some absolute criterion or 'basic reason' which is applicable to *any* practical deliberation, independent of context or horizon of self-interpretation. Taylor describes the process of practical reasoning as follows.

We show one of these comparative claims to be well founded when we show that the *move* from A to B constitutes a gain epistemically. This is something we do when we show, for instance, that we get from A to B by identifying and resolving a contradiction in A or a confusion which A relied on, or by acknowledging the importance of some factor which A screened out, or something of the sort. The argument fixes on the nature of the transition from A to B. The nerve of the rational proof consists in showing that this transition is an error-reducing one. The argument turns on rival interpretations of possible transitions from A to B, or B to A.²

The logic of practical reasoning thus has a narrative form in the specific sense that it is how we account for lived changes in moral outlook, and therefore in one's identity or 'self'. It has an *ad hominem* argumentative structure, in that it is addressed from one particular lived identity to another, as that identity is expressed in action-orientations.³ I might see myself

now, in comparison to what I was like before, as less confused about, or as more sensitive to, what really matters or is worthwhile. Practical argument, according to the 'ad hominem' model, proceeds by way of "contesting between interpretations of what I have been living"⁴, and the argument is trumped by the interpretation which, of those available, gives the most clairvoyant account of the error-reduction (or otherwise) of the transition lived through. The crucial point for Taylor's model is that what *counts* as a rationally defensible gain is not determinable independently of the actual transition. There is no appeal to criteria which are neutral with respect to the content of the moral outlooks lived through, and thus no need to invoke an ideal standard which can be brought to bear *whatever* the transition is between. If such a standard were available - in this case, of the good life, or even what is morally obligatory to do - then *any* move would be justifiable in its terms. But because any move would therefore be in principle rationally determinable, the significance of the rationality of transitions themselves is lost. Failing the availability of the neutral criterion, it can then seem as if there is no scope for rational practical deliberation. Taylor's non-criterial model for practical reasoning restores such scope, by limiting it to the passage between rival practical positions. The error-reduction from A to B or vice versa does not carry over to other positions; the validity of the transition is not transitive in the strictly logical sense. This is simply a consequence of its ad hominem structure. And of course, since the validity of the position arrived at depends on its comparative superiority to the rival party to the transition, it is always vulnerable to succession by other, more perspicuous self-interpretations relative to *it*.

From these considerations, it will appear that the whole business of a cultural theory of modernity is an exercise in practical reasoning so conceived.⁵ The form of this kind of

enquiry is to trace the origins of the hypergoods valued in modern culture, as they emerge through a transition from others which compete with them for allegiance. The modern identity can be reassured if it can defend the move from one hypergood to another, in a language of perspicuous contrast which is recognizable to both parties of the transition, as an error-reducing one. The clearest illustration of how the modern identity can, in at least *one* respect, be reassured in this manner, is provided by Taylor in his discussion of the superior rationality of modern science in comparison to its pre-modern, Renaissance counterpart. We have seen that, even for Habermas, the normative content of modernity includes the attainment of 'self-consciousness', which means, in part, the acquisition of valid knowledge about the objective world. Enlightened science removes the obstacles to self-consciousness entrenched in the pre-modern world view by divesting nature of any putative intrinsic moral significance. Habermas theorizes this shift by way of a rational reconstruction of the communicative competence in differentiating the three validity claims, but Taylor rejects the idea of a logic of development on which Habermas's model is based. So how can Taylor account for the modern achievement of 'self-consciousness' in this domain? Insofar as an understanding of the physical universe is strongly valued, and pre-modern and modern science express rival, incommensurable interpretations of this valuation, then the transition from the former to the latter can be shown to represent an epistemic gain in the following 'narrative' way.

According to Taylor's story, the practice of Renaissance science was informed by the Platonic hypergood of wisdom through attunement with the Ideas which make up the cosmos. Coming to an understanding of nature required 'participating' in this Ideal order, conceived as the reality behind the changing flux of the apparent material world.

Post-Galilean scientific practice, however, dispenses with this conception. It is no longer a function of cognizing the physical universe to be attuned to it in this Platonic sense. Scientific knowledge becomes rather a function of disengagement. Attunement is reinterpreted as a projection of human qualities onto a morally indifferent material world. Apparent flux is explained by a de-divinized or disenchanting underlying order of natural laws, according to concepts proper to the intelligibility of nature conceived as the realm of law, as opposed to, rather than a part of, the realm of meaning.

What we have here are two incompatible conceptions of scientific enquiry. They are also incommensurable because they incorporate rival norms for the understanding of physical reality. One cannot simultaneously aspire towards the attuned wisdom of the Renaissance magi and the disengaged objectivity of the modern scientist.⁶ Each is defined in contrast to the other; the reservoir of concepts and standards of explanation which make the universe intelligible as something with which the enquirer can be attuned are just what make them irrational according to the conceptual scheme proper to the intelligibility of nature as the realm of law. Which, then, is the rationally superior, if the criteria for rational understanding differ between the two practices?

Taylor's answer is that rational superiority shows itself in the practices which the disengaged science makes possible. The massive technological spin-offs of modern science requires a response from the pre-modern, and this *even if* technological control is not strongly valued as a hypergood. Given that both the pre-modern and the modern are embodied beings who are active in the world, a move in science which yields "further and more far-reaching recipes for action" is, *in this respect*, bound to constitute an epistemic

gain. But this, it seems obvious, is just what increased technological control achieves. It is important for Taylor not to assume that technological control is a neutral criterion of the rational superiority of modern science, for this is a standard not recognized by those with allegiance to its precursor. But, as Taylor observes, the efficiency of technological practices is something which, *once established*, must be accounted for by both rivals. In his paper 'Rationality', Taylor suggests that the epistemic gain in the move to the modern conception lies in the simple explanation it has for these technological spin-offs which is acceptable to both sides; it is based on a science which advances human knowledge of the physical universe.⁷ But with this formulation, Taylor looks suspiciously close to invoking a neutral, universally acceptable criterion of explanation - namely simplicity - which is just what his non-criterial model of the rationality of transitions is supposed to avoid. Matters are clearer in the more recent working paper, 'Explanation and Practical Reason'.⁸ Here Taylor more consistently sticks to the point that what lies outside the standards of the pre-modern science is the *challenge* which the success of the mechanistic science posed. Admittedly, this challenge is itself only intelligible on the assumption that there is some 'human constant' common to all enquirers, and therefore something akin to a 'criterion' operative in the structure of the argument - namely "our ability to make our way about and effect our purposes".⁹ The pre-modern Platonist can dismiss the significance of being able to make one's way about by controlling nature and thereby effecting many new purposes, but Taylor's point is that the development of new recipes for intramundane action as technological spin-offs cannot be explained on the pre-modern's own terms. This is something which can only be shown retrospectively - given modern science's technological spin-offs - in the *asymmetry* of the narratives putatively explaining either the rationality (epistemic gain) or irrationality (epistemic loss) of the transition; in the latter case, a

plausible narrative of the kind which is available to the former, which would explain the transition from Renaissance to post-Galilean science as a loss of understanding of the physical universe, is not conceivably forthcoming.¹⁰

The nerve of the rational proof here is not to invalidate the hypergood of attunement, but it *does* invalidate its association with knowledge of the physical universe, and the methods of natural science. Taylor also illustrates how the demise of a morally significant cosmos can be brought into play for demonstrating the rationality of transitions in the moral domain of a disenchanted universe. Where the cosmos is believed to instantiate a moral order, and the order is believed to be mirrored in social hierarchies, acts transgressive of the social hierarchy can be perceived to merit particularly cruel forms of punishment, irrespective of any moral principle - which is nonetheless acknowledged - to minimize pain.¹¹ But with the rational transition to the post-Galilean picture of the cosmos, one factor which contributed to the justification of overriding the principle of minimizing suffering loses its hold, and this makes a *prima facie* case for thinking that the transition to less cruel practices of punishment in the modern world has some claim to rationality. Once again, it is assumed here that there is some human 'constant' at work in the justification - namely, the acknowledgement that minimizing pain is morally desirable. While neither this nor the 'improved recipes for action' invoked in his narrative explanation of the rational transition to post-Galilean science constitute 'criteria' in the sense of being external standards applicable independently of any transition, Taylor suggests that normal moral or practical reasoning is even further away from the criterial model. In the biographical context, the case is simplified by there being only *one* person living through the transitions. As was noted above, the preferred conclusion of the practical deliberation is shown *directly* in the

perspicuity of the self-interpretation of the transition lived through; there is no need to refer "even to the differential performance of the rival views in relation to some [other] decisive consideration".¹²

In this section I have summarized Taylor's claims concerning the narrative structure of practical reason. His strategy has been to show a common structural core - captured in the idea of narrative - to the intelligibility of rational transitions in science and practical deliberation. But the outcome of the rational transition to post-Galilean science is a divorce between the attainment of valid knowledge of nature and the hypergood of attunement. So if the modern identity is to be reassured with respect to 'attunement', the source of attunement must be articulable by other means. The view that there are no other means, that there is no world with which humans can be morally attuned, that *any* conception of a moral reality is deluded - a mere projection of human attributes onto intrinsically meaningless substance - is something Taylor believes can be refuted with just the same narrative tools as Platonist science can be. Further, to refute this family of views would be to remove what Taylor claims to be an entrenched source of inarticulacy concerning the predicaments facing the modern identity; that the tension between cognition and identity is equivalent to a metaphysical gap between facts and values. In the next section, I turn to the second definitive feature of the substantive model of rationality espoused by Taylor's cultural theory, that the conclusions of practical reasoning admit of truth.

(7.2) Cultural Theory and Moral Realism

Among the tensions Taylor claims to be immanent to the modern identity is the tendency, strengthened by acultural theories of modernity, to frustrate its own articulation.¹³ Typical

of the culture-neutral mechanisms which these theories invoke, as we noted in (1.1), is the general capacity to separate pure 'facts' from pure 'values'. The separability of fact and value, Taylor suggests, is premised on a deeply ingrained principle of modern thought; that "concerning strong evaluations, there is no truth of the matter".¹⁴ So deeply ingrained, indeed, that even cultural theories are prey to it. And it is adherence to this axiom, Taylor argues, which is responsible for cultural theory's supposed affinity with relativism. In this section, I will examine the version of moral realism which Taylor proposes on the basis of the impasse faced by theories of modernity which deny the truth-evaluability of strong evaluations. First, I will look at the difficulties Taylor exposes in the 'neutralist' position associated with acultural theories. I have already raised objections of this sort against Gellner's theory (2.3), so I shall only briefly recapitulate them here. After considering why Taylor thinks that relativist versions of cultural theory covertly reproduce the source of some of these difficulties, I then turn to his favoured 'realist' version - focussing specifically on the sense in which concerning strong evaluations, there is a truth to the matter. Taylor's moral realism, we note, relies on the acceptance of what he calls the 'Best Account' principle. I then raise some objections to this principle as a resource to be exploited for 'realist' purposes.

If living within a framework of strong evaluations is not an *existential* option for human beings, neither is it a *methodological* option for the theorist of human agency to utterly dispense with the horizon within which human beings understand themselves. Taylor supports this view with the following consideration. Any culture-neutral mechanism which is invoked to explain away or 'finesse' the self-understanding of participants in modern and other societies, he argues, prejudices any account of why it is this *particular* culture which

obtains in virtue of the general mechanism. For instance, if one holds that religious beliefs are explicable in virtue of some general mechanism for social integration, one will not feel (methodologically) impelled to enquire into the *content* of the identity furnished by such beliefs. But since any such beliefs could be explained in this manner, this approach is unable to account for the specific significance of the particular manifestation of the 'culture-neutral' phenomenon. It is unable to explain why there obtain *these* details of a form of religious life - or mechanism for social integration - rather than others, since each would fit the theory equally well. But as soon as the theorist inquires into such details, he or she encounters the methodological imperative of grasping the self-understanding of the agent. And this, in general, means "being able to apply correctly the desirability characterizations which he applies in the way he applies them".¹⁵

Now the acultural theorist might accept this, without wanting to endorse the ontological commitments implied by the horizon of self-understanding. It is likely that the theorist will disagree with some of the beliefs by which the agent understands himself; so while these beliefs ought to feature in the theorist's explanandum, they ought not to determine the preferred explananda of the agent's thought and action. The way the world is, including the human world of which the agent is a part, need not be how the agent thinks it is. For the acultural theorist, according to Taylor, this means that the language of the explananda must be culture-neutral. But to be culture-neutral, it must be expunged of the evaluative biases which, given the premises of acultural theory, distort its capacity to describe the world felicitously. These biases can then be interpreted as non truth-evaluable expressions of pro- or con- attitudes. Since they are non truth-evaluable, they have no place in the explanatory language of the theory. The acultural theory, if it is to admit of truth, ought

to be neutral on such matters. But it can only be evaluatively neutral once it has identified the genuinely descriptive claims hidden in the object-language of self-understanding. The acultural theory which finds itself in this position is then impelled to transpose the object-language of self-understanding into one which has purely descriptive criteria of correct application; correct, that is, independently of culturally specific evaluative conation.

The problem with this move, Taylor notes, however, is that the criteria for the correct application of key descriptive (or classificatory) terms in the language of human self-understanding are *inextricably* evaluative. Following the argument presented in (2.3), to strip certain concepts of their evaluative force - far from 'purifying' their descriptive, truth-evaluable content - distorts this content beyond recognition.¹⁶ This means that the transposition of the language of self-understanding into a neutral, value-free vocabulary, cannot be achieved without a revision of the original vocabulary of self-understanding. The transposed language is a revision because the criteria for the correct application of its concepts are not materially equivalent to the criteria for the application of the concepts of the original object-language. But a grasp of the original criteria is a condition of the methodological requirement, mentioned above, to understand the agent. This requirement seems to be flouted by the 'neutralist' position because one cannot understand the agent without appeal to the evaluative criteria filtered out in the neutral, purely descriptive vocabulary. One is not able to apply correctly the desirability characterizations which the agent applies in the way he applies them without the linguistic resources which make such application possible. And the canonically transposed language of the acultural theorist is no such resource.

As Taylor describes it, the acultural neutralist position is motivated by the perceived need to defuse the value/ontological commitments of the desirability characterizations which define the identities of modern and pre-modern human beings. But this is exacted at the price of misconstruing the meaning of these commitments, and hence of misunderstanding the (self-interpreting) beings it aspires to explain. While cultural theory eschews both culture-neutral mechanisms and vocabularies, Taylor suggests that there is nevertheless a widely held version of it which also distorts the beings of its subject-matter by defusing the value/ontological commitments implicit in the language which expressed them. Rather than attempting to 'finesse' the language of self-understanding in the manner of aculturist theory, this version of the cultural-theoretic approach effects a similar neutralization of the value/ontological commitments of the language of self-interpretation by relativizing them to the form of life in which they have a use. Taylor calls this the 'vulgar Wittgensteinian' position.¹⁷ It involves the suspension of judgement on the evaluative force *and* the ontological commitments accompanying the self-understanding of the agent. Suspension is required because the form of life in which the self-understandings (strong evaluations) are intelligible is not the kind of thing which can either be affirmed nor denied. Since the theorist grasps the language of self-understanding of agents when she partakes in the form of life which the language expresses, and since "forms of life are incorrigible"¹⁸, so also, according to the Vulgar Wittgensteinian, are the self-understandings of the agents.

Taylor is not happy with this conclusion; "It seems to espouse", as he puts it, "an almost mind-numbing relativism".¹⁹ He suggests that the root of its inadequacy lies, like the neutralist position, in its refusal to take seriously the ontological commitments of strong evaluations. Both positions falter in adopting the thesis that concerning strong evaluations,

there is no truth of the matter. If we are to move beyond the equally inadequate alternatives of interpretative distortion or mind numbing relativism, it seems, we need to be realists about strong evaluations. Before moving on to the kind of realism which Taylor proposes as the only credible option for a theory of modernity, I need to make a preliminary reservations about the strategy followed so far. For even acultural theories which insist on a sharp break between fact and value need not make it in the way depicted by Taylor. Habermas's theory, for instance, is equally critical of the 'conative' analysis of value. As we have seen, his view is that there is indeed a 'world' to which value-judgements are apt. This is the world of interpersonal relationships about which normative claims can be correct or incorrect. But the criterion of correctness in this domain is not evaluative truth, but prescriptive validity or normative rightness. Habermas's position is thus anti-realist with respect to evaluatives, but realist with respect to prescriptives. And this means that Habermas's view isn't 'neutral' in at least one of the senses in which Taylor uses the term, for it intrinsically involves critical engagement with normative validity claims. But the separation of validity claims, even if raised simultaneously, generates its own problems; including the distorted representation of the rationality of evaluative claims considered in chapter five.

So far I have been considering the negative case put by Taylor against not taking the moral realist road. I turn now to positive claims Taylor makes for the realist interpretation of strong evaluations. Taylor's version of moral realism turns on what he calls the 'Best Account' (BA) principle. Although Taylor doesn't provide us with a precise definition of this principle, its meaning and realist thrust is most readily discernible in the following passage;

How else to determine what is real or objective, or part of the furniture of things, that by seeing what properties or entities or features our best account of things has to invoke? Our favoured ontology for the micro-constitution of the physical universe now includes quarks and several kinds of force...we have our present array of recognized entities because they are the ones invoked in what we now see as the most believable account of physical reality.

There is no reason to proceed differently in the domain of human affairs, by which I mean the domain where we deliberate about our future action, assess our own and others' character, feelings, reactions, comportments, and also attempt to understand and explain these. As a result of our discussions, reflections, arguments, challenges, and examinations, we will come to see a certain vocabulary as the most realistic and insightful for the things of this domain. What these terms pick out will be what is real to us here, and it cannot and should not be otherwise. If we cannot deliberate effectively, or understand and explain people's actions illuminatingly, without such terms as 'courage' or 'generosity', then these are real features of the world.²⁰

The real, then, is what is *disclosed* by the best account of any given domain available to us. In this respect, there is continuity between the human and the natural sciences. But as their domains differ, so will the kind of vocabulary which features in the best account. The crucial point for Taylor's realist thesis is that the best account of how human beings lead their lives cannot but incorporate a language of strong evaluation. In the previous chapter we considered Taylor's phenomenological argument for the inescapability of frameworks of qualitative contrasts. We are now in a position to consider why such a *phenomenological* account must feature in the *best* account. Given the BA principle - that the best account determines our ontological commitments - we can understand Taylor's moral realism as a doctrine of *phenomenological ontology*.

According to the orthodox view, the task of phenomenological ontology is to produce 'pure descriptions' of experience, and hence of the world as it appears from the point of view of

an engaged inquirer. According to the standard objection to it, the exclusive focus on description rids it of any explanatory power, and hence of any competence to determine ontological commitments. Taylor's BA principle is meant to disarm this objection, not by denying the relevance of explanations for phenomenological accounts, but by emphasizing the explanatory force of the language through which human beings make sense of their lives.²¹ We make sense of our lives, Taylor insists, by getting clear on the issues which face us in the living of them, and this inevitably involves using a vocabulary of desirability characterizations like 'courageous', 'honest', 'cruel', 'generous', etc. If I were deprived of this linguistic resource, then I would not understand my life so clairvoyantly, nor deliberate so effectively, nor assess the behaviour of others so competently. But this means that an explanation of the leading of lives which dispensed with these terms would also miss out on the insights they disclose, on the sense actions have, and hence on something fundamental to human agency which *needs* explaining. The BA principle dictates that the best account of human life both explains it and makes the most sense of the living of it, "and no epistemological or metaphysical considerations of a more general kind" can trump it.²² Thus Taylor can claim that there is no good reason why we should not speak of a 'moral reality', in the sense that "what you can't help having recourse to in life is real, or as near to reality as you can get a grasp at present".²³ If phenomenological reflection of the kind considered earlier shows that one cannot but have recourse to strong evaluations, then the values invoked by them are real.

I have already raised doubts about whether phenomenological reflection must show this. But irrespective of those objections, is it correct to think that we are dealing with a *uniform* sense of 'reality' here? The weight Taylor is giving to phenomenological accounts of moral

experience needs to be seen as a counterpoise to the general epistemological and metaphysical considerations mentioned above, the kind which issue in the modern construal of the fact/value distinction.²⁴ The key consideration Taylor targets is the view that truth is the exclusive property of propositions which feature in an 'absolute' account of reality; one that "prescinds from anthropocentric properties, in particular the meanings that things have for us", and which - in virtue of this - "offers the best explanation, not only of the extra-human universe (that much seems now fairly clear), but of human life as well"²⁵. Reality, according to this view, is what is disclosed to the third person perspective of the disengaged enquirer. Corresponding to this commitment is the view that "the terms of everyday life, those in which we go about living our lives, are to be relegated to the realm of mere appearance"²⁶, and hence bereft of explanatory force. This is the deep-seated premise of what Taylor calls 'naturalism'. According to the naturalist picture, the only account of reality is that articulated in absolute, non-anthropocentric terms. It follows, given this premise, that human affairs, like the physical universe, "ought to be maximally described in external, non-culture-bound terms".²⁷ But Taylor insists that the absolute account does not give us the best account of human affairs, for the reason that the meanings things have for us must feature in the understanding and explanation of them. Although the reality disclosed by the value terms featuring in our best account is 'dependent on us', it is no less real, given human existence, than the reality disclosed in the absolute account of physics.²⁸

If one means by reality what is not conditioned by human existence, then Taylor's position on science is realist, but not on morals. There is in fact a tension with respect to the supposed symmetry and asymmetry between scientific and moral realism. Taylor is (or

aspires to be) both a scientific and a moral realist insofar as he holds that there is a truth of the matter disclosed in the best account of both domains. Only the former, however, is absolute or human-independent. But what right do we have to say that the latter is 'absolute' if it depends on what gives us 'our' best account? It looks as if we can have either the best account principle, or the absolute conception of science. If we have the latter, reality has nothing to do with us, so we can't be moral realists (in Taylor's sense). If we have the former, there is no absolute reality which prescind from the human articulation of it, so we can't be scientific realists. The tension can be put in another way. 'Best accounts' are true in virtue of the reality they disclose by way of a 'perspicuous contrast'. The truth of scientific theories, on the other hand, *once the absolute conception is granted*, requires no such contrast. All it requires is correspondence. Taylor's equivocal use of the term 'reality' corresponds to these different senses of 'truth'.

The point I am trying to make can be put in still another, and I think decisive way. A major difference between best accounts of the moral and physical domain which bears on the sense in which they are true is that Taylor's moral realism allows for the possibility of incommensurability between moral truths²⁹. But this can't be the case for scientific truths, conceived according to the absolute conception, for otherwise either or both of the incommensurable truths would not be absolute. The absolute conception of reality is in principle something about which there is (potential) universal consensus. Of course, it has been a commonplace since Kuhn to talk of incommensurable scientific theories. Here we need to distinguish two kinds of incommensurability. First, there is the (let's call it 'first-order') incommensurability of theories *within* the absolute conception, say between Newtonian and Einsteinian physics. Distinct from this, there is the ('second order')

incommensurability between the absolute conception which both Newtonian and Einsteinian physics explores, and the conception of 'correspondences' explored by its pre-modern counterpart. Taylor would accept that the change from Newtonian to Einsteinian physics is of a different order to that from Renaissance to modern science. But the best account principle doesn't make sense of this difference. And this is partly because within the absolute conception, the criteria of best accounts - such as predictive power, simplicity, and technological control - are *not* retrospective. And for this reason, I suggest, Taylor would not recognize first-order incommensurability of the kind just mentioned. Why should this matter?

It matters because the Best Account principle is *tailored* to suit incommensurabilities. Its point is to address *crises* within traditions or identities. Following MacIntyre, Taylor exploits the retrospective reason of narrativity for the resolution of conflicts arising from the unavailability of a common measure. Were there to be a common measure, say for instance as there is in what Kuhn calls science in its 'normal', 'puzzle-solving' period, the BA principle would lose its point. And this is where Habermas wedges his fundamental objection to Taylor's view. For in addition to the 'world-disclosive' or 'expressive' dimension of language, there is also the 'problem-solving' capacity of language exploited in communicative action. It is in this dimension where the truths of 'perspicuous contrasts' are *tested* against the requirements of intra-mundane practice.

At the beginning of this section, I considered the objection Taylor makes against the validation of exclusively acultural theories; that they are either consistent with a wide variety of cultural practices and beliefs, and therefore non-explanatory of particular ones,

or they attempt to explain particular practices at the price of distorting their meaning. I then considered the Best Account principle as a basis for validation which supposedly favours a cultural theory construed along realist lines. In the next section I want to consider in more detail the role of practice in this kind of validation, and the interpretative model of language and critical reflection on which it is based.

(7.3) Language and the Limits of Cultural Theory

I will briefly consider three kinds of objection which might be raised against the role played by language in Taylor's cultural theory and the model of critical reflection it recommends. First, I bring Wittgensteinian objections, not too dissimilar to those raised against Habermas in (4.4), to bear on Taylor's presentation of the role played by language in the philosophical foundations of a 'cultural theory' of modernity. I will do this by focussing on a misunderstanding betrayed in the charge Taylor levels at the 'vulgar Wittgensteinian'. Second, I return to the issue of systematically distorted communication, which Habermas takes to be beyond the grasp of hermeneutic reflection, and for that reason a suitable point of departure for an alternative conception of critique. Here I take up the objection anticipated in (6.2), and to which I will return in the following chapter, that disturbances in self-formative processes from outside language may put motivational constraints on what is worth pursuing. Third, I re-introduce the case, again to feature later, for a qualified acultural theory on the basis of the second objection.

(1) The charge levelled by Taylor against the vulgar Wittgensteinian position, we saw, is that since it proposes that forms of life are incorrigible, it effectively consigns its holders to an absurd, mind-numbing relativism. But the mistake implicit in this attribution might

be thought to reveal a deeper confusion in the foundations of Taylor's model of critique. This model, as we have seen, is based on *self-interpretations*. For Taylor, human beings - beings with the identity of agents - are essentially self-interpreting. This is something we unavoidably are as language users. Without a horizon of linguistically articulable strong evaluations through which to interpret ourselves, Taylor urges, we would not be recognizably human. The task of a cultural theory is to get clear on the self-interpretations which constitute the modern identity, by way of the contrasts within it and with foreign identities. It seeks an interpretative language of perspicuous contrast:

This would be a language in which we could formulate both their way of life and ours as alternative possibilities in relation to some human constants at work in both. It would be a language in which the possible human variations would be so formulated that both our form of life and theirs could be perspicuously described as alternative such variations. Such a language of contrast might show their language of understanding to be distorted or inadequate in some respects, or it might show ours to be so (in which case, we might find that understanding them leads to an alteration of our self-understanding, and hence our form of life...); or it might show both to be so.³⁰

We saw in (7.1) how Taylor appeals to the human constant of being able to 'find one's way about' through 'effective recipes for action' when arguing for the superior rationality of modern science over its Renaissance precursor. The language of the latter was shown to be inadequate in *this* respect. Or more precisely, the contrast by which the transition to its successor could be articulated was less perspicuous than that available to its rival, for it was unable to account for why the latter should be so successful in its capacity to 'effect purposes'. The most general of human constants is the orientation to the good; inwardness, the affirmation of everyday life, and the disengaged or romantic relationship to nature represent the modern variations. But at this juncture, the Wittgensteinian can take issue

with the view that a 'language of understanding' must be a language of interpretative contrast. In (4.4), I tried to show that Habermas's thesis that action oriented to reaching understanding (the raising and redeeming of validity claims) does not stand up to the challenge of Wittgenstein's considerations on rule-following. But those considerations generate analogous difficulties to the view that understanding must involve interpretation. Just as understanding requires the grasp of something which is not a justification (or a validity claim), so what this grasp is of cannot be an interpretation - on pain of infinite regress. For interpretation involves the replacement or substitution of one sign by another, but if these signs are not to 'hang in the air', if they are to have a meaning which is the object of a *possible* conflict of interpretation, there must be a way of understanding them *independent* of and logically prior to their interpretation. For Wittgenstein, understanding is - in the first instance - "a way of grasping a rule which is *not* an *interpretation*, but which is exhibited in what we call obeying a rule and going against it in actual cases"³¹.

The point is that in presenting forms of life as things which are incorrigible or corrigible, Taylor implies that they are essentially sources of interpretation. He can then argue that interpretations are never incorrigible (this is Taylor's realism), and that *a fortiori*, neither are forms of life. If by 'form of life' is meant a cluster of self-interpretations - or in Taylor's sense a 'culture' - then the thought that each is incorrigible is 'quite 'mind-numbing'. Indeed, it is incoherent, given the logic of interpretation. But this cannot be what is meant by a form of life, since it includes the idea of an agreement in practice which conditions the possibility of interpretative dispute. Tully reminds us that interpretation is a reflective activity occasioned by a *lack* of understanding.³² It is a testament to the breakdown in pre-interpretive understanding; to sign usages which have

become problematic. But it represents only *one* way of reflecting critically with established meanings and practices, not the only, unavoidable, or essential way. The point is congruent with the objection made in the previous chapter that Taylor's depiction of an identity crisis represents not the limits of human agency as such, but of a reflective human being whose self-interpretation is of fundamental concern. This might lead us to the further conjecture that Taylor's emphasis on interpretation is itself an ethnocentric symptom of an identity whose nature has become, in modern conditions, intrinsically problematic.

(2) Back in (1.3), we noted that in his dispute with Gadamer, Habermas turned to systematically distorted communication as the paradigmatic instance of a linguistic phenomenon the grasp of which lies beyond the scope of hermeneutic reflection. Actions and utterances which appear incomprehensible both to the subject of them and to the ordinary participant in everyday discourse, Habermas argued, are rendered intelligible not in virtue of communicative competences which can be assumed in any speaker of a natural language, but rather by certain *theoretical* assumptions held by the competent observer - in this case, the psychoanalyst. Unlike hermeneutic self-interpretation, the 'scenic understanding' mediated by the theoretically informed analyst has explanatory power, in that it identifies the causal origins of the incomprehensibility in early disturbances to the agent's self-formative process. This synthesis of causal explanation and linguistic understanding gives systematically distorted communication an intelligibility it could never have under merely hermeneutic premisses.³³ However, according to the argument presented in (3.4), it is difficult to see how 'scenic understanding' could have the causal significance which Habermas claims for it. The claim to equivalence between the original, symptomatic, and transference situations on which the explanatory force of the scenic understanding rests,

carries an onerous epistemological mortgage from which this version of 'depth hermeneutics' may not recover. We also saw that the controversial epistemological claim to equivalence is adventitious to the claim to 'depth' made for psychoanalytical self-reflection, for the self-narration of the analysand can be *theoretically* informed without it requiring insight into singular and determinate causal origins. Furthermore, in the light of Taylor's conception of hermeneutics, self-interpretations can feature in *explanations* of action without subsuming the particular event under a general causal law. It seems, then, as if Habermas's appeal to systematically distorted communication doesn't significantly challenge the role ascribed to language within Taylor's cultural theory and its hermeneutic model of self-reflection.

But something important has been missed from this representation of the significance which systematically distorted communication has for Habermas. We can agree with Taylor's hermeneutic insight that the explanatory value of the theoretically informed autobiography lies not so much in its correspondence to an original fact of consciousness - something which would provide a 'criterion' for the best explanation - but rather in its comparative superiority to rival self-interpretations. We can also agree with Bernstein's insistence upon the internal relationship between the verification conditions of the general psychoanalytical interpretation of self-development and the productive, self-transformative character of the particular self-narration. But none of this touches upon Habermas's hypothesis that the disturbances in self-formative processes which issue in systematically distorted communication have their origins in *power* relations which *structure* the social context of self-formation, quite independently of the capacity of the particular agent to recognize them. To be sure, the acceptability of this hypothesis would turn on it being the best account of

those available. But any account which failed to reckon on the socializing process which conditions the individual's capacity for linguistic deliberation over identity-carrying strong evaluations must stand to it at a disadvantage. What systematically distorted communication shows is that since the *initiation* into language which conditions hermeneutic deliberation *through* it is subject to disturbance, a model of 'intact' contexts of self-formation is required even of hermeneutic theories, even though that context is not itself amenable to hermeneutic reflection. And unless such an account of an intact self-formative process which *conditions* competent desirability-characterization attribution is forthcoming, then agents engaged in practical deliberation over what is most worth pursuing may do so in self-deception - or even worse, they may be lead to radical practical error. The point becomes more pressing when directed against the general agenda of Taylor's cultural theory.

(3) The task of Taylor's cultural theory of modernity is to bring to self-clarification the sources of significance which have historically shaped the modern identity. It deploys a genealogical method by tracing these sources back to their pre-modern precursors, and in showing how they came to take such a widespread hold on people's allegiance, it helps to explain the transition to modernity. But Taylor distances himself from the 'Idealist' claim that an account of the kind he offers is sufficient for explanatory purposes.³⁴ He acknowledges that any satisfactory diachronic causal explanation would have to take account of how material, economic, social, and psychological forces shaped the conditions for cultural allegiances, and he observes that practices expressing certain aspects of the emerging modern identity were (and remain) "brutally" imposed.³⁵ But he insists that the causal path between conceptions of the good and material/economic forces is two-way. Ideas shape and can transform the 'realm of necessity', just as the realm of necessity must

take a certain form in order for conceptions of the good to be both convincing and sustainable.

But again, if a theory was available which showed a systematic, non-contingent relationship between the realm of necessity and the *space* for articulating and transmitting conceptions of the good life, then Taylor's account would stand at a disadvantage to it so far as explanatory power goes. The agenda for such a theory would be set not so much by the conflict between rival conceptions of the good, but by the social space made available for practical rational deliberation concerning them. Most notably, the significance Taylor attaches to 'ordinary life' or 'the life of production and of the family' would thereby be altered. On the new agenda, this could be interpreted as a realm within which rational practical deliberation is prone to systematic distortion, due to extra-linguistic pressures which themselves *inflect* the linguistic resources available to hermeneutic reflection. But this would be a phenomenon which lies beyond the scope of a cultural theory, even though it indirectly affects the content of that about which it seeks clarification. Whereas I earlier criticized Habermas's conception of the lifeworld for being insufficiently sensitive to the substance of normative conflict generated *internally* to it, it now seems that Taylor's concept of a background framework of strong evaluation is insufficiently sensitive to normative conflicts generated *externally* to it.

I will briefly return to this issue of how ordinary life features on the agenda of Taylor's cultural and Habermas's acultural theory of modernity in the next chapter. But before that, I must turn to another source of contention with a non-vulgar Wittgensteinian 'expressivist' position. Habermas's proceduralism is motivated on the one hand by scepticism with regard

to a 'pre-modern' evaluative realism, and on the other by suspicion of self-re-endorsing, anti-Enlightenment prejudices. If a substantive theory of practical reason need not be burdened by the insufferable mortgage of 'metaphysical premises', it might still be supportive of conservatism in morals, politics, and practices of critical reflection.

(7.4) The Expressivist Mediation of Morality and Ethical Life

We saw in section (5.3) how Habermas's deontological specification of the moral domain generated a problem of mediation between morality and ethical life. We also saw that Habermas's insistence upon the separation of these domains is motivated by the worrying anticipation of conservatism when moral reflection denies itself the possibility of transcending the parochialism of the lifeworld. I now want to consider how these issues of conservatism and mediation can be addressed from *within* an expressivist/hermeneutic framework for critical reflection. Now although not directly addressed by Taylor, this very task has been taken up by a philosopher wholly sympathetic to his expressive realist commitments - Sabina Lovibond in her book *Realism and Imagination in Ethics*.³⁶ Once the different senses of conservatism associated with these commitments have been disentangled, Lovibond urges, an expressivism purged of spurious conservative associations offers itself as a sound philosophical basis for critical reflection. Although Lovibond's representation of the Hegelian distinction between morality and ethical life is idiosyncratic, her discussion remains useful since it challenges a strong motivation for Habermas's quasi-transcendental standard of critique and procedural conception of rationality.

Let us first briefly remind ourselves of these 'expressive realist commitments'. Language is conceived as the 'expression' of a practice or 'form of life', and it is agreement in this

which ultimately conditions the use meaningful signs. Language embodies thought in a similar way to the sense in which, according to expressivism, social institutions are the embodiment of a culture. The family, civil society, and the state, for example, are considered as the 'objective' expression of the identity of those who participate in them ('objective spirit', as Hegel called them). Similarly, the objectivity of meaning is thought to consist in the publicly shared rules which determine the correct application of concepts. 'Moral reality' is the name given to the world as disclosed by particular moral concepts, and access to this world is conditioned by competence in the use of moral concepts. Moral concepts are capable of disclosing an 'objective' moral world - a world which is in some sense 'found' - in virtue of rules which put constraints on the purely subjective or merely capricious use of them. Moral understanding, or competence in the use of moral concepts, requires the acknowledgement of a public or intersubjective *authority* in the way a moral concept can be applied correctly; that is, in sentences which are 'true'. Moral objectivity, on this view, refers to a material 'pull' which exercises itself on every participant in the moral language game. Only on the basis of an 'agreement in action' does the possibility of meaningful, and hence objective, moral discourse arise. A moral reality can disclose itself, therefore, only to human beings who have been initiated into a language, which is to say to participants in a form of life which is 'expressed' in its language and other social institutions.

Very broadly speaking, this conception contrasts with an 'instrumental' understanding of language according to which word-meanings are in general determinable prior to their use; for example by mental states which are considered as logically prior to the words which name them. Similarly, it is opposed to 'contract' theories of society, according to which

the legitimacy of social institutions is modelled, in the first instance, upon a rational agreement between sovereign individuals. The expressivist view takes objection to the voluntarism implicit in both these models. It questions the coherence of the 'instrumentalist' and 'contractarian' transcendence of that embeddedness of individuals in language and society which conditions our ability to think and act as concrete individuals. Indeed, it is only in virtue of this embeddedness in language and social institutions that one counts as a rational agent at all. What counts as rational (and as 'real') is a function of the rule-governed institutions which make up the practice in which one is embedded. It is these actually existing rules, norms, institutions and customs which express a moral reality, participation in which constitutes one as a rational agent, which are called in the expressivist tradition the *Sittlichkeit*.

Is Habermas's worry is that these expressivist commitments lend themselves to conservatism in ethics and politics well founded? It is true, Lovibond submits, that the existence of a moral world requires continuity both diachronically and intersubjectively between participants in the moral community. It is in virtue of such continuity that 'spirit' is made objective, and thus that we can talk of moral objectivity at all. It is also incoherent to think that we can abstract completely from the *Sittlichkeit*, in the kind of all-at-once overthrow of prevailing meanings, pretension to which informs radical scepticism. This much, indeed, is accepted by Habermas. But just as radical scepticism is the limiting case of doubt concerning moral objectivity, so the unquestioning acceptance of consensual norms - the view that any deviation from them is "a sufficient reason to debar the person concerned from the shared way of life in which every individual must participate if he is to sustain his identity as a rational subject" ³⁷ - is the *limiting case* of moral certainty. Between these

limits of total conceptual anarchy and conservative submission to the status quo lies the space for rational criticism, the expressivist will claim. This space is opened up by the relatively loose norms which guide the applicability of most moral (essentially contestable) concepts, and by the availability of 'recessive' rationalities which challenge the orthodox expression of normativity. Further, there is no way of telling *a priori* the extent of moral certainty which is sufficient to sustain a person's identity as a rational agent, nor the degree of conformity which is sufficient to preserve the identity of a form of life. (It is up to the community "to decide when anomie has gone too far"³⁸).

The fear that expressivism leaves the individual trapped in the parochialism of the lifeworld is defused, Lovibond maintains, as soon as we distinguish between (a benign) 'transcendental' and (a harmful) 'empirical' parochialism. Transcendental parochialism refers to the inescapability of the conceptual scheme to which we are transcendently related as *corporeal* beings, that is, as beings whose efforts at rational justification rest on the natural limit of a sub-linguistic consensus in forms of life. It is the parochialism of beings whose reality is mediated by language, but whose language is not *grounded* in rational choice or agreement. This much, that rational subjectivity rests ultimately on a shared form of biological life - on something animal - cannot be changed, according to expressivism. Transcendental parochialism of this kind, however, is not in any way informative about where the limits of rationality are to be drawn with respect to any *particular* belief or action. Its point is the philosophical one of establishing the conditions of rational thought and action in general; but this does not allow the expressivist to pronounce on the rationality of any (non-philosophical) actual thought or act. Empirical parochialism, on the other hand, turns this expressivist acknowledgement of the immanence

of rationality to forms of life as a whole into a *policy* for conserving the actually existing language games, and the institutions which embody them, from change. But this conservative policy, Lovibond insists, is wholly adventitious to expressivism *per se*. She attempts to explain the conservative interpretation of expressivism by appeal to a contrast between contingent *attitudes* which may be adopted towards those who do not conform with the prevailing *Sittlichkeit*. She does this by way of a phenomenology of dissent.

The policy adopted to those who fail to find their identity expressed in the dominant communal institutions, and who in turn express their dissent in "*unsittlich* behaviour"³⁹, typically depends on whether a 'participant' or an 'observer' attitude is taken towards them, Lovibond suggests. In the former case, the challenge of the dissident to the prevailing norms of rationality is recognized as a *potential* act of rational subjectivity, or as worthy of consideration as a candidate for *Moralität*. Accordingly, the dissident is regarded as someone "to be talked *to* rather than *about*". But this attitude may be withdrawn, and replaced by an 'observer' attitude according to which the dissident's participant status in the language game is annulled. The dissident then suffers the fate of an "excommunication"; she finds herself banished, by virtue of the anomaly she expresses, from the community of rational beings. The agent of *unsittlich* behaviour, then, is considered not as someone *with* whom dialogue is appropriate, but as something to be 'managed', 'cured', or 'trained'. "The conservative", Lovibond suggests, "is prone to think that deviation from the main channel of the language game is a sufficient ground for the objectification of the deviant".⁴⁰ But this *need* not be the 'expressivist' response to those whose behaviour or beliefs fall outside the accepted criteria of rationality. Faced with a moral anomaly, one has the option of taking the performative or the objectifying stance towards the agent of it:

the destruction of the dissident is brought about by the actions of individual persons who could, in principle, choose to do something other than what they currently do: and this understanding will lead us to anticipate the comment that the apparent penchant of dissenters for self-mutilation may be no more than a reflection of the violence emanating from within the moral organism.⁴¹

This is the violence, Lovibond remarks, of a wholly static moral organism, one in which the only recognized moral obligations are those embodied in the moral reality as currently constituted. The objectification of the dissident, in other words, is the policy which accords with the conception of moral objectivity espoused by 'empirical' parochialism. To take the option of adopting a performative attitude, on the other hand, is to acknowledge the vulnerability of the moral organism to rational change. But the rationality of this change is not *initially* expressible in an 'objective' form; it must rather be understood in its 'subjective' aspect, as carried by the 'imagination'. The task then becomes one of reintegrating the subjective and the objective perspectives on moral reality, which "must result from the consciously willed establishment of an expressive relationship between ourselves and our public institutions".⁴²

Hence a problem of mediation or 'reconciliation' arises for the expressivist. At what point can we reasonably claim that the inner, subjective perspective on the moral world (*qua Moralität*) carried by the imagination is reconciled or mediated with the outer, objective, or public perspective on moral reality (*qua Sittlichkeit*)? Since the obstacle to this reconciliation is the 'empirical' parochialism of the historically given *Sittlichkeit*, resistance to which motivates rational dissent, the problem can be reformulated as how critical reflection can be pushed back to the point of purely transcendental parochialism. The

possibility of identifying empirical parochialisms, according to Lovibond, rests on the availability of concrete conceptions of a less partial or arbitrary moral world, as expressed in the lives of those participating in (or imagining) some other community. But acceptance of this, Lovibond conjectures, presupposes commitment to a limiting point at which the ground to be covered in the overcoming of empirical parochialism runs out. Such an absolute conception of moral reality would be "one in which the individual or local perspectives of all human beings would find harmonious expression".⁴³ At this 'absolute' point, parochialism would be solely transcendental in kind, and the form of life which expressed it would be in agreement, as Lovibond puts it, with "universal reason".⁴⁴

From a substantive political point of view, there is much for Habermas to be in agreement with here. Lovibond's proposed 'absolute' conception of morality as what is embodied in a form of life in agreement with 'universal reason' by virtue of the local perspectives of *all* human beings finding harmonious expression, clearly bears the hallmark of the utopian moment of Habermas's counterfactually presupposed communication community. And Lovibond's attack on the conservative expressivists surreptitious presumption of a static moral organism converges with the space Habermas opens up for reflexivity and fluidity within the lifeworld when reproduced by communicatively rational action. But once the expressivist has recourse to a distinction between empirical and transcendental parochialism, she need not feel imprisoned by the resources of the actually existing ethical customs from which, Habermas fears, expressivism as such offers no escape. Given this distinction, Habermas's worry can be seen as directed towards a *policy* for excluding participant status to the moral anomaly represented by the dissident, but this policy can now be seen as adventitious to the philosophical mainspring of the expressivist conception of language itself.

(7.5) Conclusion

But while the utopian, critical direction of expressivism and Habermas's theory may be aligned, there is a reversal at the level of justification. According to Lovibond's view, 'universal reason' is still a 'parochialism' of sorts - the parochialism of embodied linguistic, self-interpreting animals - and is conceived as a limiting point of a process of reflection which begins with the content of actual linguistic practices. The gap between this point of departure and the limiting conception of a concrete universalism, for Lovibond, is covered by the *imagination* of the critic. But for Habermas, the process of justification begins with idealized presuppositions which *precede* concrete conceptions of their embodiment. According to Habermas's conception, we begin by abstracting ourselves into the idealized universalism of communicative reason, and then cover the ground *backwards* to the parochialism of the actuality of the *Sittlichkeit*. But this is an otiose move for the expressivist, since the question of what reason there is for recognizing the objective claims of *moralität* cannot be answered in advance by appeal to necessarily presupposed structures of thought or action, since it cannot be decided in advance of the availability of *moral sources* to empower morally motivated action. I have already considered the difficulties of mediation to which Habermas's recommended exercise in abstraction leads. But the problem of mediation has by no means vanished with the adoption of an expressivist standpoint. For what now needs to be reconciled are the subjective perspective on the moral world as carried by the imagination and oriented to moral sources, and the objective or publicly instituted ethical life. The question remains open as to how, *under conditions of modernity*, what Lovibond calls an expressive relation can be willed between ourselves and our public institutions. For this, in Taylor's terms, would require that moral sources assume an objective form in the 'public order of references'. But is it really conceivable

that in the modern world an expressive relation can be established with an objectively articulable moral order? And if not, what are the implications for the *form* the imagination must take if it is to gain access to moral sources?

CHAPTER EIGHT: THE CRITIQUE OF REASON: INSTRUMENTAL, FUNCTIONAL OR PROCEDURAL?

(8.0) Introduction

By reference to the work of Habermas and Taylor, I have been examining two ways in which philosophical insight into a tension between cognition and identity can clarify the idea that modernity - construed as a horizon of thought and action or as a mode of being - is disposed to crisis. Both Habermas and Taylor undertake to put the intuition that modernity suffers from a self-undermining dominance of instrumental reason on a philosophically sound basis. Further, they share the view that this basis needs to be recovered from conceptual distortions built into the 'subject-centred' paradigm of reason. According to that paradigm, rationality is conceived in terms of a cognitive/instrumental relation between a subject and an object. The modern distinction between subject and object, both authors stress, needs to be seen in contrast to the pre-modern fusion of cognition and identity in cosmological world views. Coming to knowledge of the cosmological order, according to these views, is co-extensional with finding one's place in the moral order. But with the modern turn, an object of cognition becomes defined as something which is *external* to human identity. This makes room for an 'instrumental' paradigm of rationality, based upon the external, antagonistic, and deeply paradoxical relation of a subject to an object.¹ Both Habermas and Taylor offer accounts of how Post-Enlightenment thought struggles to re-establish an internal relation between cognition and identity, without returning to the kind of 'substantive' paradigm of reason embodied in the pre-modern conception of cosmological orders. But the hold of subject-centred reason is such, both our authors agree, that even the most radical critics of instrumentalism can unwittingly reproduce it. For this reason,

Taylor concurs with Habermas's thesis that the 'total critique of reason' in contemporary neo-Nietzscheanism founders precisely in virtue of the hidden subject-centredness which anchors it.² In other words, both claim that this kind of critique commits a performative contradiction; it cannot make sense of its own normative presuppositions without appealing to those very norms which are, or are entailed by, the object of its critique.

In (8.1) I consider the extent to which, from Taylor's critical perspective on instrumentalism, Habermas's theory is *also* guilty of this charge of 'performative contradiction'. For Habermas, we have seen, the critique of instrumental reason is best seen as a redemption of the claims of linguistically constituted intersubjectivity from encroachment by a derivative (and coercive) functional reason. An internal relation between cognition and identity is regained, we might interpret Habermas as claiming, in the communicative *conditions* of self-formation. But by making this into a claim about the *procedural* nature of communicative rationality, Habermas is open to the charge that the normative presuppositions of communicative action, explored in 'discourse ethics', are irrelevant for questions of identity. For *that* there are such presuppositions in no way informs me why it is good for me not to contradict them, unless I am already committed to an identity which has allegiance to the good of consistently following the communicative procedure.³ But whether I should have this commitment - a question of identity - is not itself answerable by reference to the procedure. Since the procedural paradigm of practical reason cannot make sense of the norm presupposed in the motivation to follow the procedure, it is guilty of performative contradiction. For Taylor, *any* procedural account of rationality reproduces the basic errors and blindnesses of instrumentalism, amongst which is the denial of substantive goods (a conception of freedom and dignity) presupposed in that

denial. But if, as I argued in chapter five, Habermas's *general* critique of instrumental reason is *not* procedural - because it is not independent of a substantive conception of human dignity - then what remains of the charge that Habermas's model of critique is (implicitly) subject-centred? After rejecting various formulations of this charge as resting on misunderstandings of Habermas's position, I turn (8.2) to the claims Taylor makes concerning artistic modernism. Contrary to Habermas's subjectivist interpretation (as an unfettered exploration of human subjectivity), Taylor reads modernist art as striving after 'epiphanies' which disclose a real, trans-subjective moral order, but one only accessible through a language of personal resonance. While these claims might be valid, I go on to question their compatibility with the claims about identity and practical reasoning which featured in the transcendental arguments for moral realism considered in the previous two chapters.

In (8.3), I consider the implications of Taylor's claim that the moral sources disclosed by modernist epiphanies can be tapped *only* by being indexed to a personal subjective vision. Taylor's suggestion that moral sources can no longer structure a public order of references is, I argue, problematic both in terms of his own account of how legitimation crises emerge in modern societies, and in terms of how these crises may in principle be resolved. I suggest that his bleak diagnosis results from an arbitrary generalization of the fate of a theistic moral source, a fate which again reminds us of the young Hegel's diagnosis of the positivity of religion. This puts us in a position to reassess the relevance of the model of an ethical totality which Habermas draws from Hegel. I propose that taken as a model of self-formative processes, the Hegelian view as interpreted by Habermas can at least obviate some of the unnecessary difficulties generated by a theistic commitment adventitious to

Taylor's basic hermeneutic insights. One of these insights is not only that artistic modernism is sensitive to the experiential consequences of instrumentalism - an insight which, with its focus on intersubjectivity, Habermas's approach elides - but also that it engages in a form of critical practice which is the most appropriate response to those consequences. And this gives the cutting edge to Taylor's conception of philosophical critique; to recover contact with moral sources beyond the atomized self which are not only occluded in modern forms of inarticulacy, but realizable only in virtue of the modern creative imagination. In (8.4), I compare the responses of Habermas and Taylor to the critical capacity of artistic modernism. I suggest that these different responses correspond to the roles of what I call the objective and the subjective critic. In my conclusion, I will sketch the different perspectives brought by the two kinds of critic to an issue at the heart of the political problem of modernity's self-reassurance; ecology. Using Habermas's distinction between the problem-solving and world-disclosive dimensions of language, I state a promissory case for both kinds of criticism, under the condition that the needs to which they are addressed are conceptually and practically distinguished. I conclude by proposing this as a policy for the critique of instrumentalism in general.

(8.1) Procedural Reason and Performative Self-Contradiction

Critics of modernity repeatedly identify the dehumanizing effects of a form of existence rationalized along predominantly 'instrumental' lines. Philosophical critics have located these effects in very general structures of thought and action which have been presumed to constitute the rationality of the modern horizon of self-understanding. The critique of 'Instrumental Reason' traces these structures back to the fundamental relationship between a subject and an object, and is informed by something like the following picture of modern

rationality. As a structure of thought, 'Reason' is the subject's capacity for correctly ordering and accurately representing external neutral objects, a capacity which requires the minimization of subjective value or meaning-projections onto the objective world. Classical empiricist philosophy can be seen as an articulation of this horizon, and criticism of modern instrumentalism often goes by way of an attack on the philosophy which reinforces this picture of human beings and their relation to the world.⁴ Likewise, classical utilitarianism can be seen as drawing upon and perpetuating this picture of human beings and their dealings with the world.⁵ As a horizon of human agency, objects are still intrinsically neutral, but are invested with value insofar as they are instrumental to the purposes of the human subject. But practical reason is incapable of divining what these purposes should be, for there is no objective moral ordering of purposes to be divined. There is, however, one purpose which, even as an object, the subject must consider in its practical deliberations; that of self-preservation. Self-preservation can thus be construed as the only end which is rationally determinable - the only point at which practical reasoning extends beyond the province of calculating efficient means to subjectively chosen ends. Cognition in a disenchanted world is of objects accurately represented by a subject which, in its capacity to represent itself, can control and manipulate objects *and* itself for the purpose of its self-preservation. But by turning itself into an object, by determining the province of practical reasoning as that of the calculable and manipulable, the subject loses the aura of its *own* intrinsic meaningfulness, and hence the point of its practical deliberations. The critics of Instrumental Reason further argue that the dominant practices and social institutions of modernity - the market, capitalism, and bureaucratic administration - are attributable to the overriding hold on western civilization of the imperatives of self-preservation; where the notion of self-preservation is modelled on the primordial

relation between a subject and an object. Modernity *as a whole* is diagnosed as being geared towards the same kind of unconstrained (and ultimately self-destructive) expansion and control as is prefigured in the subject-object relation.⁶

While Habermas and Taylor agree that the phenomena identified by the critique of Instrumental Reason are real enough, they both resist attributing them to the domination of 'Reason' as such. That the critique of Reason should take this path betrays the presupposition that rationality must be understood *either* in its Platonic, substantive form, *or* as modelled on a means-ends/representative relation of a subject to an object. The critique of Reason, if it is to be capable of articulating that by appeal to which it is critical, must show why this opposition is inadequate. Habermas does this by way of proposing an alternative paradigm of communicative reason. The subject-object relation can then be regarded as parasitic upon a more primordial context of communicative intersubjectivity, from which the norms for the critique of instrumental reason, now considered as a critique of "an unleashed functionalist reason of systems maintenance", can be derived.⁷ So the crisis which the critique of Instrumental Reason diagnoses as a dialectic of the subject-object relation, is diagnosed by the critique of 'Functionalist Reason' as a disfigurement of intersubjectivity. For Habermas, the subject-object relation is introduced as an alien element "into relationships that by nature follow the structure of mutual understanding among subjects".⁸ And this division and usurpation of the rational potential of communicative action corresponds to the one-sided rationalization (along the aspect of cognitive/instrumental mastery) of the lifeworld promoted by the unfettered, functionally self-preserving imperatives of the economic and administrative sub-systems. The critique of Functionalist Reason avoids the normative impasse consequent upon the assumption made

by the critique of Instrumental Reason that;

the disenchantment of religious-metaphysical world views robs rationality, along with the contents of tradition, of all substantive connotations and thereby strips it of its power to have a structure-forming influence on the lifeworld beyond the purposive-rational organization of means. As opposed to this, I would like to insist that, despite its purely procedural character as disburdened of all religious and metaphysical mortgages, communicative reason is directly implicated in social life-processes insofar as acts of mutual understanding take on the role of a mechanism for coordinating action.⁹

The question I now want to raise is whether Habermas's metacritique of the critique of Instrumental Reason - that it cannot make sense of its normative foundations - can itself be turned against the critique of Functionalist Reason, insofar as the latter appeals to a conception of reason with a "purely procedural character". Taylor gives a critique of procedural reason which effectively poses this challenge. Matters will be clearer if we distinguish Taylor's general objections against proceduralism from his specific criticisms of Habermas's version of it. This is because the nature of the proceduralism attacked differs in the two cases.

There are two quite different contexts in which Taylor examines the nature and shortcomings of a procedural conception of reason. In the first place, he associates it with the paradigm of rational disengagement inaugurated by Descartes but only fully developed in Locke's conception of a 'punctual self'. "Disengagement is always correlative to an 'objectification'", Taylor writes, where "objectifying a given domain involves depriving it of its normative force for us."¹⁰ The mechanization of the scientific world picture involved such an objectification; it neutralized an order of 'Ideas' or 'correspondences' to a norm-free mechanism. While Taylor acknowledges that this transition represents an

epistemic gain so far as our understanding of the physical universe goes, he claims that the adoption of a disengaged objectifying stance towards the subject of experience itself is more a source of bewitchment and error. It is this radicalization of disengaged reason, Taylor suggests, which issues in a model of the punctual self. The procedure of disassembling and reconstructing items of experience according to fixed canons of thinking - which is definitive of rationality for the disengaged inquirer,¹¹ - objectifies the pre-disengaged flow of experience by depriving it of any normative epistemic content; it is homogenized to the level of cognitively innocent and impersonal inner sensations which have a scientifically describable external cause but no genuine object. Taylor points out that much of human experience and knowledge is either unintelligible or arbitrarily dismissed along these lines, but it is the ethical rather than the epistemological motives behind proceduralism which really interest him.¹² For by adopting an objectifying stance towards the self, the prospect opens up for unprecedented control and self-remaking for which procedural reason is the instrument. Such is Locke's 'punctual self', which Taylor defines as the subject who takes a radical stance of disengagement to himself;

To take this stance is to identify oneself with the power to objectify and remake, and by this act to distance oneself from all the particular features which are objects of potential change. What we are essentially is none of the latter, but what finds itself capable of fixing them and working on them.¹³

Reason is construed as the procedure to be followed in exercising this control. Following it is the responsibility of each individual agent, who in taking the radically reflexive objectifying stance towards his beliefs, habits and desires - in short, what is pre-given in his embodiment - can remould them with the dignity afforded to the autonomous, atomic, knowing subject. As Taylor remarks, the validity of the first-person perspective is

paradoxically undermined by the very disengaged stance it is a condition of.¹⁴

If this is what it means for rationality to be procedural, clearly communicative rationality does not have a purely procedural character. On the contrary, the foregoing depiction of 'procedural reason' is almost identical to the portrayal of the very 'subject-centred reason' in opposition to which Habermas offers his communicative model. First, in defining the objectification perpetrated by the disengaged self in terms of norm-neutralization, and associating this with social practices (or technologies of the self) of discipline and control,¹⁵ Taylor is putting just the point expressed in Habermas's critique of the norm-free mediation of subjects determined by the functional mechanism of administrative power (the objectifications of functional reason). Second, it is just to avoid the paradoxical objectifying self-relation of the subject that Habermas seeks to ground rationality in the linguistic medium of intersubjectivity. Nor is the disengaged proceduralism identified by Taylor to be conflated with Habermas's rational reconstructions. The latter, unlike the former, involve a recapitulation of already existing but pre-theoretical competences, not their objectification into a norm-free, wilfully manipulable mechanism.¹⁶ As we have seen, it is when practical reason is exclusively modelled on a rational reconstruction that problems arise within Habermas's account. Procedural ethics provides us with the second context for Taylor's general critique, but so far he has not shown why a procedural conception of rationality is necessarily tied to an 'instrumental' conception of it.

Taylor's general attack on procedural ethics focusses on its refusal to acknowledge qualitative contrasts. In typical fashion, he responds to this refusal by way of a genealogy of what motivates the espousal of moral proceduralism. First, there is the family of

epistemological and metaphysical presuppositions behind the 'naturalist' temper; that the human sciences be modelled upon and continuous with the natural sciences. Accordingly, like objective states of affairs "human affairs ought to be maximally described in external, non culture-bound terms".¹⁷ But in addition to this, Taylor suggests that the proceduralist ethic is bolstered by a cluster of unavowed moral motives. First on this hidden agenda is affiliation to the affirmation of everyday life, and a suspicion of any notion of the 'higher'. The rejection of qualitative distinctions can be seen as a liberation from the crushing demands of the 'higher', or as a modest antidote to the smug self-satisfaction felt by those who claim to meet them. Second, there is an implicit commitment to a modern conception of freedom, for example in contract theorists who insist that "normative orders must originate in the will".¹⁸ Third, there is the emphasis on practical benevolence and an associated presumption that intellectual inquiry ought to be directed towards relieving "the condition of mankind". This motive can be compounded with a suspicion that theorists focussing on qualitative distinctions are self-indulgent or socially irresponsible; unconcerned about justice, altruism, and the like. Fourth, the proceduralist ethic seeks to avoid "parochial ethical principles", which Taylor attributes to affiliation to the hypergood of 'purity'.¹⁹ This affiliation also disposes the proceduralist to make hard distinctions between purely moral and other values.²⁰

Taylor's point in making the above objections is to identify goods which are presupposed in the procedural account of practical reason; especially the goods of freedom and benevolence as they are supported by the hypergoods of rational disengagement and the affirmation of everyday life. And only if one is motivated by these hypergoods, will there be any reason to abide by the ethic; to follow obligations determined by the rules of justice.

As soon as one accepts these goods, then they must be accounted for in a conception of practical reason. *Given* the good of freedom, purity, and the like, *then* one can argue about procedures. But a theory which dismisses the kind of goods which feature in qualitative contrasts will be inarticulate about its normative foundation. Any proceduralist ethic cannot avoid this performative contradiction, so Taylor argues.

How applicable are these general criticisms to Habermas's theory? What first needs to be stressed is that one could accept the kind of theory Habermas proposes without denying the validity of qualitative distinctions. Habermas recognizes the importance of strong evaluations, it is just that he places them outside the domain of morality and pure practical reason. I have already argued (5.3), using reasons similar to those invoked by Taylor, why this taxonomy is both unfortunate and misleading. This led me to propose that Habermas's theory as a whole - that is, as addressing the problem of modernity's self-reassurance - is best interpreted in 'substantive' rather than 'procedural' terms. The applicability of the general critique against proceduralism to Habermas's theory can be made clear by considering Taylor's objection that concerning the 'communicative ethic';

The crucial point would not be the presuppositions of communicative action, but rather the way in which human identity is formed through dialogue and recognition. The argument would turn on the acknowledgement, in other words, of a crucial human good, and not simply on the pragmatic contradiction involved in the violation of certain norms.²¹

Insofar as the rational reconstructions of discourse ethics goes, this objection holds. But the wider aim of the theory of communicative action is to establish just this thesis; that *pathologies* ensue when individual and social identity is formed otherwise. Where Taylor

goes astray is in identifying the 'procedural' nature of communicative rationality with procedural practical reason. If Habermas were to modify his theory in the way proposed in chapter five, it can establish that there are human goods of autonomy and solidarity which can only be realized under conditions of undistorted intersubjectivity. The pragmatic contradiction which is relevant to Taylor's point is more the paradox identified by Weber, which Habermas theorizes as the violation of lifeworld norms by the system. Communicative reason is precisely not procedural in Taylor's sense because it is internally related to the goods which it conditions as a structure of self-formative processes.

The crucial difference between Habermas and Taylor lies in what they take to be most structurally significant in the process of self-formation. This is most evident in the role they give to 'everyday life'. For Taylor, everyday life - the life of production and the family - is of significance in virtue of the motivational capacity of the strong evaluation which affirms it as a hypergood. This valuation, Taylor suggests, features as a hidden support for the procedural ethic. But the significance of everyday life in Habermas's theory is quite different. For not only is everyday life the resource of life goods which may be affirmed, it is also the *context* of the self-formative process. This context is not itself a strong evaluation, but a condition of it. On the one hand, the life of production and biological reproduction (institutionalized in the family or not) is not something about which the social animal *homo sapiens* has a choice, and therefore is not something we can understand primarily as an affirmation. On the other hand, the dominant *mode* of production and biological reproduction clearly structures the self-formative process independently of whether the life of production and the family is itself strongly valued, or indeed of any strong evaluations. As Habermas insists, the socialization which everyday

life structures is also an individuation; the two cannot in principle be separated.

There is something misleading, then, in construing everyday life as an object of strong evaluation. The issue here is not, as Taylor suggests, suspicion of any notion of the 'higher'. It is rather that the everyday life of production and biological reproduction is a locus of power.²² The suspicion is more that notions of the 'higher' often mask the real sources of this power, both material and *simultaneously* symbolic or 'cultural'. Taylor would not deny this obvious point, it is just that his approach to everyday life directs us away from putting it. The motive behind Habermas's theory need not be so sinister as to make us suspect cultural theories of the kind adopted by Taylor of social irresponsibility. It is better to think of the former as highlighting the political question of power which the latter puts out of focus. The former approach can focus on the self-formative process which the power relations embodied in modern forms of production and biological reproduction disturb, whereas the latter focusses on these phenomena as strongly valued and hence an *expression* of identity.

Habermas's 'proceduralism', we have seen, boils down to his commitment that *self-formative processes can be explained by rational reconstructions*. Discourse ethics is but a species of this genus, and by focussing on the species, Taylor distorts Habermas's intent in ascribing to communicative reason a 'purely procedural character'. But as we have already seen, Habermas is also committed to the view that disturbances in the self-formative process can be understood by narratively carried out self-critique. And here lies the rub of Taylor's objection that Habermas's communicative ethic cannot make sense of its normative foundation. For in order to motivate action, rational reconstructions remain to

be internalized into an identity which only a life narrative can provide. As Habermas is aware, rational reconstructions *per se* cannot empower subjects to continue their disturbed self-formative process; they cannot address the motivational problem of 'why change?'. But this is a necessary condition of successful critique; there must be a passion to undertake it. The communicative ethic, and the critical powers of the theory of communicative action, is only coherent given the acknowledgement of this passion. Habermas might be correct to insist that the violation of communicative norms by the strategic actor is not an option from the point of view of the lifeworld, but it is only not an optional *reason for acting* for the agent who has *internalized the theoretical perspective of the lifeworld* into a narrative perspective on his or her own identity as constituted by an ethical totality which awaits acknowledgement. Once acknowledged, Taylor's attack on Habermas the proceduralist no longer has a target.

(8.2) The Epiphanic Mediation of Cognition and Identity

I have been arguing that, in certain crucial respects, Habermas's theory is not of the 'proceduralist' kind which is attacked by Taylor, and that consequently there is a danger of overstating the differences between their critiques of instrumental reason. Habermas's communicative reason is not 'procedural' in at least some of the senses Taylor gives to the term, because it is *internally* related to human constants which can without distortion be called 'the good', while the conception of reason which informs Taylor's critique of proceduralism is not 'substantive' in Habermas's sense because it does not appeal to a Platonic, cosmological moral order. Both recognize that the tension between cognition and identity which is built into subject-centred reason cannot be overcome by either assimilating identity to cognition (the absolutist or Raving Platonist position) or by assimilating cognition

to identity (the relativist, neo-Nietzschean position). But deep differences remain, of which the following two are the most significant. First, Taylor takes Habermas's account to be cripplingly anthropocentric, since it leaves no room for goods which are not internally related to human capacities or needs. Second, Taylor argues that the exploration of such goods, while unintelligible on Habermas's account, is in fact central to the phenomenon of modernist art. Since, as Habermas accepts, this phenomenon informs the problem of modernity's self-reassurance, Taylor can claim that Habermas's misinterpretation of modernism has a distorting effect on his formulation of the self-reassurance problem. In short, Taylor takes the view that Habermas is insufficiently radical in his critique of subject-centredness, that this is most evident in his interpretation of artistic modernism, and that the correct interpretation points to something which might redeem the modern identity from the subject-centredness or instrumentalism which has led it to crisis. The prospect of a non-anthropocentric ethic that it opens up I will consider later, for it presupposes the validity of claims Taylor makes about artistic modernism.

Taylor's difference with Habermas on the significance of artistic modernism goes to the heart of their conflicting interpretations of the problem of modernity's self-reassurance for the following reason. The significance of art in the critique of Enlightenment, for Habermas, lies primarily in its potential as a surrogate for religion as a bonding force between socialized individuals.²³ In Taylor's terms, Habermas focusses on the public consequences of the decline of religion, that is, on the demise of religion as a medium of sociation, and hence on the gap it leaves with respect to the sources of solidarity required for social integration. But besides the public consequences of instrumentalism - the rectification of which critics of Enlightenment have sought in aesthetically transformed civic

religion - there are also experiential consequences to which artistic modernism must be understood as a reaction. Both Taylor and Habermas identify the experience of objectification; of the reification of subjectivity wrought in instrumental, capitalist, modern society. But whereas Habermas discerns the explorations of and experiments in 'pure' or 'unfettered' subjectivity as the principle of modernist art - and celebrates it as an important antidote to the phenomenon of reification - Taylor interprets the subjectivism of modern art as a *route* to something greater, rather than as the destination of the modernist creative moment as such. Again, though both Habermas and Taylor want to redeem the claims to validity of modernist works of art from the positivist view that science holds exclusive rights on truth, they differ on what they take to be the nature of this claim. While Habermas stresses the claim to authenticity - to genuine expression of subjectivity - in the modernist art-work, for Taylor this view represents a massive oversight regarding the task of the modernist artist, one which betrays the vestigial grip of subject-centred reason on Habermas's theory. This is because of the modernist aspiration towards, and exploration of an order in which the artist is set as a locus of moral sources. This "search for moral sources *outside* the subject through languages which resonate *within* him or her, the grasping of an order which is inseparably indexed to a personal vision",²⁴ issues in the 'epiphanic' dimension of modernist art. The modernist exploration of a moral order through a language of personal resonance, Taylor points out, is unintelligible within Habermas's framework of modern understanding; it is a matter neither of objective scientific knowledge, nor procedural practical reasoning, nor unfettered expression of subjectivity. The validity of epiphanic art falls between the differentiated claims of objective truth, normative rightness, and subjective authenticity. But the epiphanic work of art, Taylor insists, is nevertheless criticizable; it can succeed or fail *as an epiphany*." In

this section I will now consider in more detail (1) the nature of the epiphanic claim and (2) its intelligibility within Taylor's *own* realist hermeneutic framework. In the next section, I turn to its implications for a possible resolution of the problem of self-reassurance.

(1) *The Epiphanic Claim of the Modernist Aesthetic*; Taylor's identification of and focus upon what he calls the epiphanic dimension of modernist art is in the following fundamental respect merely a continuation of his expressivist conception of language. To think of a work of art as an epiphany is to consider it "the locus of a manifestation which brings us into the presence of something which is otherwise inaccessible".²⁵ What the epiphanic work manifests "is somehow inseparable from its embodiment",²⁶ such that one cannot understand what it is "qua epiphany by pointing to some independently available object described or referent. What the work reveals has to be read in it."²⁷ As we saw in (7.4), expressivists take the same relation to hold between language and being generally. They take the view that language discloses a world to which there is no non-linguistic access²⁸; that what language reveals about the world has to be read in language, and not, say, in some sensory pre-linguistic touchstone with the world. Similarly, they deny the view that thoughts are prior to and logically independent of their linguistic manifestation, and insist rather that meanings are constitutively what they are as made manifest or embodied in language. The meaning of the epiphanic art-work is likewise indissolubly embedded in the medium of its manifestation; the visions of the epiphanic poet "give us reality in a medium which can't be separated from them".²⁹

But whereas the objectivity of meaning requires public rules for the correct application of the language, the meaning of the epiphanic art-work does not admit of such criteria. This

is because *what* the epiphany discloses - namely, moral sources - are, on Taylor's view, not publicly accessible under modern conditions. "In the Post-Enlightenment world", Taylor writes, "the epiphanic power of words cannot be treated as a fact about the order of things which holds unmediated by the works of the creative imagination."³⁰ The creative imagination, rather than the public order of meanings, becomes the indispensable locus of moral sources. Since moral sources are of "the highest moral significance", the artist can become accredited with 'paradigmatic' human status.³¹ Although Taylor gives a complex genealogy of the modern conception of art as epiphany, with its associated exaltation of the artist and the powers of creative imagination, the underlying idea here is that the world-disclosive powers of ordinary language can become impoverished. When language becomes a mere instrument for everyday, routinized interaction - as it does in the context of instrumental society - contact with moral sources becomes problematic.³² The creative imagination of the poet is then needed not only to rekindle contact with what gives human life its deeper, more fulfilling meaning, but also to discover new sources of spiritual fulfilment which are defined or completed *in* the creative revelation of them.³³

But contact with moral sources becomes problematic in other ways, evident in how Taylor traces the development of the modernist understanding of epiphany from its Romantic precursor. The Romantics, on Taylor's reading, excelled in 'epiphanies of Being'. The aim of the Romantic epiphany is to show some reality (disclosed in the work) to be an expression of something (say, the voice of nature, unsullied inner impulse) which is an unambiguously good moral source (contact with which empowered unambiguously good action). With the modernist turn, the *way* in which epiphanies make something manifest moves away from the expressive capacity of the elements of the work, while *what* is made

manifest - a moral source - becomes much more ambiguous. The first shift involves greater sensitivity to the indirect nature of the epiphany, and subsequently a need for what Taylor calls 'subtler languages' of articulation. Consequently, the notion of a 'self' - or more accurately of a 'source of the self' - which is *expressed* becomes problematized. The second shift, in addition, questions the *affirmability* of what is disclosed either by self-expression or some other form of manifestation. I will briefly consider each of these shifts in turn.

Of modernist forms of epiphany, Taylor distinguishes what he terms the 'framing epiphany', the 'intertemporal epiphany', and the 'Post-Schopenhauerian epiphany'. *Qua* epiphanies they reveal something which lies beyond the self, but unlike the Romantic epiphanies of Being, the epiphanic function is not performed by words or objects with an expressive capacity. Framing and intertemporal epiphanies frame a space between objects, or between an event and its recurrence, and it is in virtue of this space between things, rather than of the expressive power *in* them, that the epiphany can work by "bringing something close which would otherwise be infinitely remote".³⁴ This move requires a sharper reflexive focus on the particularities of language, in order to redeem its disclosive power from its degeneration into an "inert instrument" for dealing "more effectively with things".³⁵ And crucially for our purposes, Taylor acknowledges that the shift involves a 'decentring' of the self from its position as a centre *either* of disengagement *or* expressive unity. The modernist "epiphanic centre of gravity begins to be displaced from the self to the flow of experience"³⁶, such that we are taken "beyond the self as usually understood, to a fragmentation of experience which calls our ordinary notions of identity into question...or beyond that to a new kind of unity".³⁷ While the intertemporal epiphany challenges the

basic modes of time consciousness and narrativity associated with both disengaged reason and Romanticism, the modernist critique of the unitary self is most evident in the 'Post-Schopenhauerian' epiphany. Taylor continues;

The ideals of disengaged reason and of Romantic fulfilment both rely in different ways on a notion of the unitary self. The first requires a tight centre of control which dominates experience and is capable of constructing the orders of reason by which we can direct thought and life. The second sees the originally divided self come to unity in the alignment of sensibility and reason...to the degree that we adopt a post-Schopenhauerian vision of inner nature, the liberation of inner experience can seem to require that we step outside the circle of the single, unitary identity, and that we open ourselves to the flux which moves beyond the scope of control or integration.³⁸

In the post-Schopenhauerian (modernist) world "there is no single construal of experience which one can cleave to exclusively without disaster or impoverishment".³⁹ Taylor's favourite exemplar of this kind of epiphany - Thomas Mann's *The Magic Mountain* - "presents two radically incombable modes of time consciousness, one which approaches timelessness and another which is constituted by the calendar of real events".⁴⁰ Moreover, Mann shows that this "slippage from our normal sense of measured time is the essential condition for a deeper experience which opens another dimension of life".⁴¹ This recognition that human life is irreducibly multilevelled "has to be won against the presumptions of the unified self, controlling or expressive".⁴² And in this recognition we are carried "quite outside the modes of narration [common to the disengagement and expressivism] which endorse a life of continuity or growth with one biography or across generations".⁴³

The manner in which the modernist epiphany renders us close to a moral source beyond the

self thus involves a transfiguration of what is disclosed *in* the disclosure of it as refracted through the creative imagination. Now this transfigurative quality of the modernist epiphany is crucial for understanding the second rift with Romantic 'epiphanies of Being'; that concerning *what* is disclosed and its affirmability. For what the post-Schopenhauerian epiphany seems to make manifest is an enticing, meaning-endowing, but *amoral* force. Contact with it is understood as enhancing the richness of experience, as a necessary feature of a fulfilled life, even though it is not a source of goodness.⁴⁴ This breach between the claims of morality and those of the modernist creative imagination furthers the transfigurative imperative of the epiphanic art. The modernist imagination questions the goodness of being; it challenges any notion of a pre-established harmony between natural spontaneity and the claims of morality, especially benevolence. For those committed to the latter, a question of the affirmability of what the modernist imagination discloses arises: Faced with such a crisis of affirmation - quite alien to the Romantic epiphanies of Being - a radicalized transmutation of what is revealed in its revelation becomes all the more imperative. Taylor suggests that the "recovery may have to take the form of of a transformation of our stance towards the world and the self", such that, "the world's being good may now be seen as not entirely independent of our seeing it and showing it as good". He concludes that the "key to a recovery from crisis may thus consist in our being able to 'see that it is good'".⁴⁵

But such a conception of transfiguration can be developed in different ways, which Taylor illustrates by contrasting the positions of Dostoyevsky and Nietzsche. Whereas Dostoyevsky emphasizes the miraculous transfiguration of evil and degradation into an object of love through the acceptance of (others' and God's) love, Nietzsche advocates a

heroic ethic of yea-saying which not only foregoes any reconciliation with morality, but bypasses the moral consciousness and the call of its conscience to benevolence. Both advocate "a transfiguration of the world: a vision which doesn't alter any of its contents but the meaning of the whole",⁴⁶ where this new, deeper meaning, - the moral source - is brought about in the changed stance towards the self and the world. The transfiguration itself can then function as a moral source. For Nietzsche, it is the very human power to transfigure which is ultimately affirmed. But Nietzsche's idea, Taylor remarks, is vulnerable to subject-centred misinterpretation. This is the erroneous thought that just because the epiphany is indissolubly indexed to a personal vision or subjective stance, it follows that the epiphanic is *merely* the expression of subjectivity. On the contrary, the point of an epiphany is "to avoid the merely subjective". Taylor insists that "there are forms of subject-centredness which don't consist in talking directly about the self", and that overcoming subject-centredness "is a major task, both moral and aesthetic, of our time".⁴⁷ Nietzsche seeks to achieve this by celebrating "the deep recesses beyond or below the subject, or the subject's uncanny powers".⁴⁸ While acknowledging the power of Nietzsche's epiphany, it is clearly Dostoyevsky's vision of a self and world transfigured - good because 'seen as good' - in the acceptance of the love of God which resonates with Taylor. But what sense can be made of a moral *order* which is so disclosed?

The significance Taylor gives to epiphanies is partly due to the fact that such a moral order is no longer, and cannot conceivably be again, a publicly established one. For the moral sources revealed by modernist epiphanies to become a publicly established order of moral references would require that they be part of 'the tacit background of objects of reliance' of the kind described by Wittgenstein in *On Certainty*. These are matters more fundamental

than beliefs, because relying upon them conditions the possibility of genuine doubt and hence what we normally call belief. "Virtually nothing in the domain of mythology, metaphysics, or theology stands in this fashion as publicly available background today", Taylor concedes. And this means that the epiphanic disclosure of these domains - for the purpose of forging contact with moral sources - must be understood "in default of being a move against a firm background"; any such moral order indexed to a personal vision "can never become again an invoking of public references, short of an almost unimaginable return - some might say 'regression'- to a new age of faith."⁴⁹

Now it may be recalled that it just such Wittgensteinian considerations which inform Habermas's concept of the lifeworld as that context of ethical life the fragmentation and reification of which motivates the problem of modernity's self-reassurance. We have already seen (6.2) how the problem of reassurance looks when articulated by the basic concepts of Taylor's cultural theory of modernity. Before considering the implications of Taylor's claims regarding the epiphanic validity of modernist art for tackling that problem, I want to examine whether these claims are consistent with those basic concepts of identity, practical reasoning, and the good discussed in chapter six.

(2) *Modernist Epiphany and the Claims of Taylor's Cultural Theory*; I have quoted extensively Taylor's remarks on the epiphanic claim of modernism to emphasize the tension of their standing with the foundations of his proposed cultural theory of modernity. The striking suspicion is that the expressivist conception of the self undermined by modernism appears to be precisely the one exploited by Taylor in his arguments for the inescapability of strong evaluations and the narrative form of practical reasoning. The former thesis was

supported by the counterfactual claim that a life unintelligible against *some* framework of qualitative contrast would not be recognizably human. The phenomenon of an identity crisis allegedly testifies to the inescapability of the *question* of identity, defined as a stance on qualitative contrasts. But the validity of this question is *itself* put into question by the 'decentring' movement of modernism. Taylor endorses the modernist attack on expressivist notions of identity, but the question of identity is of the force Taylor requires of it only if that notion is presupposed. My identity is supposedly expressed in the stance I take, but the very adoption of a stance reflects an *already accepted* notion of what it is to have an identity. And it is doubtful whether this notion will be acceptable to the post-Schopenhauerian. For contact with 'moral sources' is for him or her not a matter of self-interpretation, and hence not something for which horizons of strong evaluation are relevant. What is relevant, even on Taylor's account, is that reservoir of amoral force which is greater than my capacities for self-interpretation.

This point raises the difficult issue of what it is in virtue of which any epiphanic vision is acceptable. Just how difficult it is can be seen by considering the tension between the claims about modernism and those about the narrative structure of practical reasoning. As we saw in (7.1), the latter turned on a phenomenology of moral growth; determinable only in biographical form. But the modernist epiphanies, in Taylor's own words cited above, 'carry us quite outside the modes of narration which endorse a life of continuity or growth with one biography or across generations'. It seems that the content of the modernist epiphany clashes with the claims made about the self as a subject of moral growth. Admittedly, Taylor suggests that we might then be taken to a higher form of unity, but Taylor is unable to clarify what this might mean.⁵⁰

Taylor might reply that lack of clarity should not be seen as an objection here, since we are dealing with epiphanies which by their nature are only articulable in a language of subjective, personal resonance. While I think this quietist response has its proper place, the problem is that *if* it applies to the 'criteria' of acceptability of epiphanies, it should *also* apply to the 'criteria' for the determination of moral growth. Since Taylor does offer an account of the latter, he should in consistency be in a position to defuse the anti-quietist objection about epiphanies. Alternatively, if he retains his quietism about epiphanies, it seems that for the same reason, he would have to forfeit a substantively rational account of moral growth. But if he takes this second option, he effectively cuts off the branch which supports not only his version of the narrative structure of practical reason and much of his moral realism, but by implication, also the *cognitive* status of epiphanies.⁵¹ "Goods", Taylor asserts, "can't be demonstrated to someone who really is impervious to them".⁵² Conversely, argument over goods always has an ad hominem dimension, in that it involves a disputed transition "from one's interlocutor's position to one's own via some error-reducing moves, such as the clearing up of a confusion, the resolving of a contradiction, or the frank acknowledgement of what really does impinge".⁵³ Since the meaning of modernist epiphanies is to disclose moral sources, this form of argument should also be able to determine the relative superiority of epiphanies. But it is difficult to see how the relative truth-disclosure of different *kinds* of epiphany can be determined in such a manner. To be sure, we think some artistic visions to be "deeper, more revealing, truer than others", and that just "what these judgements are based on is very hard to say".⁵⁴ But my point is that - in the light of the foregoing account - these judgements will be based on *different* things according to the different kinds of vision. Further, I see no reason why the same diversity, and difficulty in the wake of it, should not also apply to judgements of

strong evaluations.

To make aesthetic judgements in the way Taylor prescribes is in fact to do nothing other than invoke the Best Account (BA) principle. The best account, as we have seen, is the one which issues in a transition of error-reduction in the manner just described. But here again Taylor appears to be inconsistent in his employment of it. In (7.1) we saw how it is used to defend those ordinary, pre-scientific conceptions of personal identity and practical reasoning from reductionist revision. But now it seems that these common sense notions are quite capable of being challenged by the visions of the epiphanic poet; a challenge which, to succeed where reductionism fails, must fit into our Best Account. Yet as I have just indicated, it is very difficult to see how the BA principle can justify this claim.

Indeed, Taylor recognizes that ultimately we have to concede that "The epiphanic is genuinely mysterious".⁵⁵ Besides carrying the implication that the BA principle is itself based on a mystery, something which may be true but which fits ill with its role as the guiding tool for explaining rational scientific change, this view also implies something very awkward for Taylor's notion of human agency. Taylor continues that the epiphanic "possibly contains the key -or a key - to what it is to be human".⁵⁶ But it is difficult to reconcile this thought, however laudable, with the concept of the human agent considered in chapter six. For there we saw that the lack of a unified, integrated identity is a *pathological* human condition. Yet it is just the aspiration to overcome such an identity which Taylor discerns in much modernist epiphanic art. So Taylor seems to be accrediting paradigmatic human status to beings who may well be in the very crisis state which he earlier consigned to a pathology.⁵⁷ The post-Schopenhauerian who affirms the destruction

of a unitary identity without striving for an alternative source of unity or continuity is hence paradoxically both a paradigm of the human and outwith the bounds of the recognizably human. His or her life is both exemplary and pathological.

The paradox is easily avoided, as I suggested earlier, by taking crisis states as signifying the limit of human agency rather than its collapse into pathology. This construal fits well with the recognition of the *transfigurative* moment Taylor discerns in face of the crisis of affirmation. It is in face of the crisis that the transfigurative power of love emerges. But this crisis in turn raises a paradox over the supposed nature of moral sources. Moral sources are defined as that the contact with which empowers us to be good. But so defined, they clearly cannot generate the crisis of which Taylor speaks. Furthermore, a crisis of affirmation of which he does speak well precedes the post-Schopenhauerian turn, since it requires only a recognition of the problem of evil; on what basis can being be affirmed if it contains so much suffering?⁵⁸

There is another problem. If moral sources are as vital for the good life as Taylor claims, and the Enlightenment suppression of them is a self-mutilation, isn't their exile from the public order of references also devastating for the human identity? The question can be put another way. Can moral sources, with all their transfigurative power, escape the fate which modernity has meant for theism?

(8.3) The Causality of Fate Again

But with this question, we are back at the point in Habermas's reconstruction of the philosophical discourse of modernity where Hegel introduces his idea of the dialectic of

moral life as a causality of fate. This dialectic, as we saw in (3.1), involves a transfigurative 'seeing good' of the context of life which is the common basis of existence for self and other. Contact with moral sources, on Hegel's model, involves the acknowledgement and recognition of real networks of dependency; the networks in which the self is immersed as part of an ethical totality. This ethical totality avenges itself in the *crisis of affirmation* of the individual with respect to the *actual* conditions of self-fulfilment. We can then understand the crisis of affirmation as issuing from intersubjective conditions of self-formation in which real networks of dependency appear as a hostile fate. Hence the need for transfiguration; not as a resigned acknowledgement of the inner goodness of the actual, but as an affirmation of an occluded ethical totality in which a reconciliation between acknowledged networks of dependency and self-fulfilment becomes possible.

I want to suggest that such thoughts can help fill a lacuna in Taylor's diagnosis of the crises facing the modern identity. Besides the issue of instrumentalism, Taylor identifies two other roots of the modern moral predicament. One is that underlying the almost universally recognized standards of justice and benevolence, there is profound conflict and uncertainty over the constitutive goods behind these norms; that is, over what it is that makes human beings worthy of just or benevolent action.⁵⁹ The other is that living up to the high standards of universal justice, self-determination, and benevolence, tends to have a crippling effect on the aspiration to expressive fulfilment. "High standards need strong sources"⁶⁰, Taylor proclaims, and he goes on to suggest that only God can really provide such strength. Gone, however, are the days when The Moral Source could be part of a public order of references - failing an 'unimaginable' return to an age of faith. The fate of modernity is such that action in accord with the high standards to which modernity aspires can be

empowered only by a subjectively accessible moral source. Cognition and identity are in tension because under modern conditions the deepest source of identity has no public or *objective* foundation. The Moral Source is no longer an object of public knowledge; rather its foundation lies in the transfigurative moment of the acceptance of God's love, a moment which 'makes good' by 'seeing good', as of the kind articulated in the most resonant epiphanies of modernism.

The main weakness with this kind of diagnosis, I suggested, lies in the assumptions it carries concerning self-formative processes. On the one hand, one might object - in the manner of Rorty - that Taylor's view is insufficiently sensitive to the multifarious and mundane *accidents* of this process. Taylor's realism requires an account of why the epiphanic art work bears the capacity to resonate not only with *my* personally indexed vision, but also with that of *competent* others. Ultimately, Taylor takes the quietist option that the epiphanic phenomenon is genuinely mysterious. But it is an open question whether the phenomenological considerations invoked by Taylor are just as compatible with the view that the epiphanic is but an epiphenomenon of the contingencies of self-formation. On the other hand, one could object - more in the manner of Habermas - that Taylor's view is insufficiently sensitive to the mundane *necessities* of self-formation. By limiting himself to the 'cultural' sources of the modern identity, traceable by a genealogy of values, the *materially constrained* socializing processes which *make possible* the individuation of personal visions is all but side-stepped. The objection here would go that Taylor exaggerates the significance of phenomena like artistic modernism in this process. But the issue I want to focus on is that art can only take the significance Taylor attributes to it *if* it is embedded in the intersubjective milieu out of which subjectivities emerge. That is, if

it is *institutionalized in and significantly structures the public order of references.*

Now this implies a potentially devastating objection to Taylor's diagnostic thesis. It is crucial to Taylor's argument that societies suffer legitimation crises when their public institutions fail to express the identities of their members. Modern instrumental society fits this diagnosis not only because its dominant institutions embody values of control, discipline, and disengagement alien to and undermining of the values recognized by many of its individuals, but also because, under the pressures towards growth and atomization, they render 'norm-free' what previously was strongly valued. But how can these public consequences of instrumentalism be avoided if not by way of a re-establishment of a public order of references which embody constitutive goods? It would seem to require, *given* Taylor's preference for a theisticly interpreted Moral Source, a civic religion precisely of the kind he thinks is now unavailable. This would leave us in just that position of 'self-mutilation' of which Taylor complains in those who deny the reality of constitutive goods, for that something the love of which empowers good action would then be constitutively absent from the very intersubjective 'context of life' which conditions individuated subjectivity. Must we accept it?

If we think of moral sources theisticly as The Moral Source, then I think we must. There are three outstanding reasons for thinking this. First there is modern consciousness of time, change, and contingency which *generates* the problem of self-reassurance. Second, it is difficult to imagine that a *single* public order of references will satisfy the moral needs of moderns. And third, there is the pressure on the public order of references to meet needs *other* than moral ones. Besides the world-disclosive function of the public order of

references, there is - as Habermas insists - its requirement for *intramundane* problem-solving. Precisely because of the rapidity of change and growing complexity of these problems, the significance of this latter role seems bound to be heightened. Since civic religion is otiose in this respect, and incompatible with fundamental moral needs in the others, its fate looks sealed.

But it doesn't follow from this that a public order as a locus of moral sources beyond the atomized self must share the same fate. This becomes clear as soon as we take Taylor's suggestion that the key to a recovery from the modern crisis of affirmation may lie in our capacity to "see that it is good" neither in the spirit of Dostoyevsky nor Nietzsche - but in the young Hegel's. The transfiguration to which Hegel refers is that pertaining to an ethical totality, occluded *qua* ethical insofar as the public order of references fails to function as a moral source, actually internalized into the identity of individuals. But the nature of transfiguration is such that it is unable to specify the particular contents of the ethical totality (public order of moral references); it is the whole which is seen differently, not the parts. The working of the causality of fate is such that requires acknowledgement of that dependence on others which avenges itself on the separation invoked by the abstract particularity of the agent of subject-centred reason.

Hegel's model thus suggests a certain kind of critical practice as the most appropriate response to the crisis of affirmation. For Habermas, we have seen, the causality of fate operates as a form of distorted self-formative process, and the transfigurative love which disenchant it allows that process to continue. He then attempts to build this insight into a quasi-empirical theory of the social pathologies of modernity by way of rational

reconstructions. But because what counts as moral growth or development is determined by the substantial picture of maturity, health, or the good which provides its 'terminus'⁶¹, it is not amenable of neutral, disengaged 'empirical' questioning. In what sense are they open to questioning in a manner that expresses Hegel's insight concerning the crisis of affirmation?

(8.4) The Objective and the Subjective Critic

The philosophical critique of instrumentalism - theorized as subject-centred, functional, or procedural reason - characteristically aligns itself with modernist artistic practices. For Habermas, the imaginative exploration of 'unfettered subjectivity', the reinterpretation of needs in the light of such exploration, and the 'fluidification' of rigidified identities, are all salutary correctives to the objectification of the *subjective* world as it appears from the intersubjective horizon of a reified *lifeworld*. The lifeworld is always the focus of interest for Habermas's model of critique. The task of the critic is to diagnose identity crises in the intersubjective *mediation* of identity. Habermas recognizes that of the symptoms which accrue from cultural rationalization, the response of much modern art is exemplary. Of these symptoms, Habermas focusses on the splitting off of expert cultures from everyday life, and the differentiation of discourses and value spheres which thematize the different validity claims. The interpreter of the lifeworld can learn from artistic efforts to bridge the gap between expert culture and everyday life, as well as from works which metaphorically intermesh the validity claims. Nevertheless, the philosophical critic must respect the cognitive gains which the modern differentiation of the validity claims makes possible. Differentiation must be maintained to avoid the 'aestheticizing' of morality and politics of which neo-Nietzscheanism falls foul. But art is not an apt model of criticism for the further

reason that it neglects the 'problem-solving', action-coordinative function of language. This problem-solving capacity is built into what Habermas takes to be the primary mode of language use; action coordination on the basis of a rationally motivated agreement. The counterfactual presuppositions of communicative action, for Habermas, condition the objectivity of meaning. In prioritizing the public or intersubjective conditions of rational action, and in seeking to trace sources of discontent to the distortion of such conditions, Habermas's is a form of *objective* criticism. The world-disclosive capacity of language, on the other hand, is enhanced when the action-coordinative warrant of illocutions is *bracketed*.

Taylor's model of philosophical critique aligns itself with modernist art in an even more fundamental way. A cultural theory of modernity, we have already seen, attempts to articulate in a language of most perspicuous contrast competing horizons of qualitative distinction which inform conceptions of the good. Once one takes the ontological commitments of these horizons seriously (moral realism), the task can be seen as one of exploring the moral order in which we are set with an aim to defining moral sources. But the contemporary critic lives in an age in which a publicly accessible cosmic order of meanings is an impossibility. It follows then, that the *language* of philosophical criticism - through which the moral order in which the critic is set is made perspicuous - can exact contact with moral sources only in virtue of its 'personal resonance' with the critic. According to Taylor's conception, philosophical criticism is successful when it enables us to see things more clearly than we did before, but in the case of moral sources, this perspicuity must be one which realizes contact with the source it explores - it must resonate with us. Philosophical criticism of the kind practised and advocated by Taylor thus assumes an ineluctably *subjective* character. And for the following reason, this means that the

prospect for self-reassurance converges in a peculiar manner with the province of modernist art.

To say that a moral source resonates with us, is to express the thought that morality stands in an *internal* relationship to our identity. The person in contact with a moral source is an agent *motivated* to act morally. Gone, however, are the days when knowledge of the objective world could attune us to moral sources, and hence motivate moral action. From the pre-modern horizon, to know the world objectively is to love it; it is to be motivated by that love to moral action which is simultaneously empowered by the knowledge of the objective order of cosmic meanings. The philosophical question which haunts modernity is that which follows the break between cognition and identity; if there are no objective meanings in the world to be known, on what basis can I be reassured about the norms which fill out my identity? Three possible responses are rejected by Taylor. First, the return to Platonic moral cosmology or strong moral realism; this fails to account for the success of disengaged science and technology. Second, the affirmation of moral relativism and the powers of unleashed subjectivity; this fails to account for the phenomenological pull of values to something beyond human subjectivity. Third, the deontological/liberal compromise; this fails to account for the motivation to follow rule-governed constraints. The only credible response must be one which is compatible with the cognitive achievements of modern science, while sufficiently realistic to deal with the phenomenology of moral experience, and sufficiently subjective to account for morally motivated action.

Such a response is embodied, Taylor suggests, in modernist epiphanic art. The epiphanic work puts us in contact with the moral order in which we are set; it empowers us in the

same way as the cosmic order once empowered pre-moderns. But whereas attunement with the latter was constitutive of objective knowledge for which the locus was Reason, resonance with the former is now a matter of *subjective* knowledge for which the locus is the creative imagination. To know is still to love - and hence cognition is reconciled with identity - but in the case of the epiphanic work, this knowledge is inseparably indexed to a personal vision. Philosophical criticism as the exploration of the moral order in which we are set now has to understand itself likewise. Since the question of self-reassurance demands such an exploration, it must be addressed in a manner exemplified by artistic modernism.

The point of philosophical critique would then be to put us (moderns) back in touch with moral sources. For Taylor, we have already seen, the goal of philosophical criticism is to clarify the moral predicaments besetting the modern identity, in a language which makes perspicuous the contrast between both sides of the transition between identities. Moreover, Taylor contends, this perspicuity *empowers* me to lead the moral life in virtue of its resonance within my particular subjectivity. In what is essentially a recovery job of the same kind which is undertaken in modernist art, philosophical critique aims to retrieve contact with the moral sources occluded and suppressed in routinized and impoverished modern 'instrumentalist' thought and practices, and thereby help to *make manifest* those sources. Philosophical critique so understood is oriented towards transfigured subjectivity, and is a consequence of the recognition that a certain kind of moral objectivity is no longer defensible. "As our public traditions of family, ecology, even polis are undermined or swept away", Taylor writes, "we need new languages of personal resonance to make crucial human goods alive for us again".⁶²

Of course, I recognize that the distinction between subjective and objective criticism I am introducing here stands in need of further clarification and support. I will offer something by way of this by considering how the distinction might be *applied* to one particularly pressing problem facing contemporary modernity. It is a problem, I believe, which provides a sobering context for assessing the viability of our alternatives to 'subject-centred reason'. In my conclusion, I will indicate why, with respect to ecology, the task of self-reassurance is something in which the objective and the subjective critic must collaborate.

CONCLUSION: THE POLITICS OF DEMOCRATIC PROBLEM-SOLVING AND WORLD-DISCLOSURE

I have been arguing that Habermas's communication-theoretic paradigm for the critique of instrumentalism fails to make available crucial conceptual resources for addressing the problem of modernity's self-reassurance. As we saw, this is a point which Taylor attempts to establish by way of reconstructing an epiphanic dimension in artistic modernism. The intersubjective constitution of the subject through language in no way detracts, Taylor insists, from the validity of explorations of a moral order to which there is no intersubjective *access*. The critique of instrumentalism as undertaken by modernist art is supposed to disclose non-human goods through the personal, human resonance carried *in* their disclosure. The acknowledgement of such goods, Taylor asserts, is a key moral challenge of our time. It is a challenge which can only be met by the invention of new languages which keep 'non-human goods alive for us' - in the manner in which they are acknowledged in much modernist art. I identified this as the task of the subjective critic. But in helping to keep non-human goods alive for us, subjective criticism clearly takes on a political character. One implication is that in making non-human goods matter, the interpretative language of the subjective critic can help to motivate ecologically attuned action: "It would greatly help in staving off ecological disaster", Taylor states, "if we could recover a sense of the demand that our natural surrounding and wilderness make on us" - demands which can be voiced not in the anthropocentric language of Habermas's theory, but rather in that of a subjectively resonant epiphanic poem or philosophical prose.¹

Anything approaching a proper philosophical analysis of the claim that the natural

environment places moral demands on humans is of course way beyond my present remit. But I do feel bound, by way of a conclusion, to suggest how the two philosophical models for a critique of instrumentalism I have been exploring might be applied concretely, and I shall do this by considering in brief the light they might throw on our understanding of the political problem of self-reassurance nurtured by the awareness of environmental devastation in the modern world. The 'environmental crisis' is simultaneously a legitimation and an identity crisis for moderns. How adequately do our conceptions of philosophical critique articulate it?

At first sight, the prospect for discourse ethics in this respect seems, as Habermas concedes, rather bleak. He accepts that the "anthropocentric way of looking at things" built into discourse ethics cannot make sense of moral intuitions made manifest in compassion for "the pain caused by the destruction of biotypes".² His response is to delimit the domain of philosophical ethics; "to clarify the universal core of our moral intuitions and thereby to refute value skepticism". "By singling out a procedure of decision making", discourse ethics "seeks to make room for those involved, who must then find answers on their own to the moral-practical issues that come to them, or are imposed upon them, with objective historical force".³ But like most self-limiting manoeuvres, Habermas's looks suspiciously *ad hoc*. Nothing is explained about the *specific* intuited wrong of biotype destruction by making it external to the imputed universal core of morality. Similarly, from the point of view of an opponent who has that intuition of wrong, Habermas's response might simply seem to beg the question of what constitutes the core of what are properly called 'moral' intuitions.⁴ Neither is it self-evident that the intuited wrong of biotype destruction is a manifestation of *compassion*. It at least remains an issue whether the intuited demands that

the natural environment might make on us are based on something more than the suffering of sentient.⁵

This is the kind of issue, I want to suggest, about which subjective criticism is pertinent. The *content* of the moral-practical issues which come to the participants in the dialogical decision-making procedure depends partly on what *matters* to the participants themselves. And what counts as mattering depends upon the moral world which is disclosed to them. The task of the subjective critic is to *sensitize* us to dimensions of an intersubjectively *non-actualized* moral reality. This activity of sensitization - the creative imaginative undertakings of the subjective critic - corresponds to what might be called the politics of *world-disclosure*. In the case of ecological politics, it would involve the creation of a new language in which the natural environment is disclosed as a world that mattered independently of human interest in it, but which could show itself as such only *in* being disclosed with subjective resonance in human beings. Subjectively resonant explorations of a moral order in which we are set, of the kind achieved in in much modernist art, could render "demands from beyond the self more palpable and real for us", Taylor suggests, instancing those demands that "underlie a more-than-anthropocentric ecological policy".⁶ The politics of world-disclosure thus borrows from the form of artistic modernism for the purpose of identifying the absence of human resonance from the actual (objective/intersubjective) moral world.

To a certain extent, Habermas seems to concur with this point when he ascribes 'green' protest to sensitivity to those industrial developments which "noticeably affect the organic foundations of the life-world and make one drastically aware of standards of livability, of

inflexible limits to the deprivation of sensual/aesthetic background needs".⁷ Although within Habermas's conceptual framework it cannot be construed otherwise, this view *need* not be interpreted in anthropocentric terms. Something can only matter to a human being if, in some sense, it satisfies a need. But it doesn't follow from this that the satisfaction of a need is *why* something matters. Individual need satisfaction is a condition of the recognition of a good, but not necessarily that in virtue of which it is a good. The politics of world-disclosure concerns just this recognition; the rendering articulate of actually unrecognized or muted moral claims. If we understand 'moral' in Taylor's substantive sense - rather than the procedural sense given to it by discourse ethics - then it covers the claims which are muted *in* what Habermas calls the attack on the organic foundations of the lifeworld. One task of the subjective critic would be to make this attack more visible, to make us even *more* drastically conscious of criteria of livability, where what it is to live a fully human life might also cover the acknowledgement of non-sentient claims made vocal *through* the overcoming of aesthetic deprivation. If we think of inflexible limits to the deprivation of sensual/aesthetic needs as signifying something about the state of the moral world through subjectivity, rather than signifying something about subjectivity through the state of the objective world, then the anthropomorphic locus of world-disclosure can be combined with intuitions about the non-anthropocentric content of what is disclosed.

Political ecology, then, has this world-disclosive character. It seeks to transform, as we might say, the 'mattering' of the world. To this end it attempts to show the inadequacy of discourses which represent the natural environment as relating to humans as a 'standing resource',⁸ by way of a contrasting language which has a deeper subjective resonance. Only *given* this resonance or sensitivity can green issues matter in a non-instrumental way; it

must be *prior* to any decision-making process about human dealings with the environment. There is something in this kind of will-formation which Taylor's realist focus on qualitative contrasts captures more satisfactorily than Habermas's attempt to ground philosophical critique on the idea of unconstrained intersubjectivity or *democratic* will-formation.

Once the decision-making process has begun, however, we move beyond the orbit of world-disclosure to the politics of *problem-solving*. The critical labour of deconstructing the 'world picture',⁹ or of disclosing a world of deeper personal resonance, is essentially a preliminary task; problems might be redefined, reprioritized, or rediscovered; the relationship of the individual to the whole might be transfigured, but particular intramundane problems of action co-ordination remain. The task of the objective critic is to identify those public or intersubjective constraints on what counts as normatively valid problem-solving practices. *At this level*, Habermas's discourse ethics offers itself as a plausible model for how individuals aspiring to autonomy might have that autonomy respected simultaneously to co-ordinating their actions in a rational manner. For the institutionalization of the presuppositions of communicative action, to the degree of structuring the public order of references, does not suffer from the difficulties facing the conception of nature as a moral source *under conditions of modernity*. Leaving aside the strong claims Habermas makes for discourse ethics - that it articulates the common core to all moralities, and reconstructs the essence of strictly 'moral' intuitions - is it so crippling anthropocentric as to be of *no* relevance for an understanding of the environmental crisis which motivates political ecology?

Although practical problems have a meaning which is partly determined by the best

available language of self-interpretation, they are also imposed on agents, as Habermas reminds us above, by an 'objective historical force'. This is the force not only - in Habermas's vocabulary - of the symbolic a priori of an intersubjectively shared lifeworld, but of the material a priori of the state/economic system. To focus exclusively on the politics of world disclosure, in the manner of subjective critique, is to open oneself to the 'Idealist' charge that one is tacitly presupposing that a particular horizon of world interpretation is the *cause* of the environmental crisis. It is to neglect, that is, action consequences the meaning of which is graspable not from the point of view of the lifeworld, but from the vantage point available to the theorist of the historical forces arising from the self-perpetuating mechanism of the modern/capitalist state/economic system. Since it is not graspable from the point of view of the lifeworld, the lifeworld itself is a misplaced object of critique. The crucial point is that the proper object of critique would be the *lack* of accountability of action-consequences to decision making procedures circumscribed by the normative content of the modern lifeworld, rather than actions which are themselves thought to be expressions or manifestations of that lifeworld. This is not necessarily to say that the source of accountability for environmentally destructive action consequences is the procedure of decision-making - as Habermas is forced into advocating - but it is to give due focus on the *de facto* threat to the environment consequent upon the bypassing of communicatively reached solutions to problems of action co-ordination.

Of course, there can be no guarantees in this domain. It is logically possible that communicative problem-solvers will continue that 'rage against Being' wrought by the systemic imperatives of profit and control. But it seems hardly unwarranted to claim that processes of genuine (non-systemically distorted) democratic will-formation are *less likely*

to generate self-undermining conditions of environmental devastation. Though not a guarantee, three rather obvious considerations can give us confidence that democratic problem-solving - theorized along Habermasian lines as action co-ordination governed by the rational principles of a communicative ethic - would turn the tide against environmental destruction, and hence support the view that discourse ethics is of *some* relevance for grasping the philosophical basis of a normatively secure political ecology. First, and most obvious, the control exercised by communicative actors over decision-making processes concerning technology is not to be confused with unfettered technological control over nature; a point which Habermas has struggled to make plain since his earliest writings. Second, since only norms which lead to the satisfaction of *universalizable* interests pass through the legitimating communicative procedure, then maxims which when institutionalized frustrate the satisfaction of the interest in a healthy environment *which everyone has* will not do so. Here, indeed, the problem of reaching a rational consensus over incommensurable interests which in other contexts beleaguers communicative ethics does not arise.¹⁰ Third, the *interference* of democratic problem-solving processes by the profit imperative *at the level of production* gives an eminently plausible *explanation* of what it is which provokes the need for critique.¹¹ And as yet, we do not know what it would be like for this critique not to be needed.

To be sure, even if we acknowledge that the environmental consequences of communicative action are less likely to be as destructive as those of actions mediated by money and power, this still fails to account for the 'green' intuition, mentioned by Taylor, that mute creatures are intrinsically worth preserving; that 'Being' is worth letting be. Two rejoinders can be made here. First, no guarantees can be offered by the subjective critic either. The latter

can attempt to disclose a believable world in a personally resonant language in which Being is an apt object of reverence, but the person so affected cannot escape from the nexus of action-coordinative problems that need to be solved under the material constraints of a social world which, if genuinely structured by communicative action, would accommodate *different* personally resonant visions of the natural world. There is an in-built indeterminacy to both the politics of world-disclosure and of democratic problem-solving.

Partly because of this indeterminacy, the critique of instrumentalism must tap the resources of both objective and subjective criticism. I have referred to the task of the subjective critic as that of presenting a horizon of self-interpretation which discloses a believable and desirable world with a particular kind of significance. And this leads me to the second rejoinder; that the believability and desirability of worlds disclosed through subjectivity is partly conditioned by their *uncoerced* resonance. If this latter point is neglected; the awkward question will loom as to what distinguishes the subjective critic from the propagandist. Taylor acknowledges the danger of such a slippage. On the task of articulating moral sources, he notes that "the whole thing may be counterfeited", and continues:

This is not to say that words of power themselves may be counterfeit. But the act by which their pronouncing releases force can be rhetorically imitated, either to feed our self-conceit or for even more sinister purposes, such as the defence of a discreditable status quo. Trite formulae may combine with the historical sham to weave a cocoon of moral assurance around us which actually insulates us from the energy of true moral sources. And there is worse: the release of power can be hideously caricatured to enhance the energy of evil, as at Nuremburg.¹²

The awkward question is this; how are we to tell the difference between the wholesome act

of pronouncement and the rhetorical imitation; between the caricature and the real thing; between the counterfeit and the genuine article? Not, of course, simply by the degree of subjective resonance - as the powers of the propagandist or charismatic dictator amply testify. As a subjective critic, Taylor can only pit the power of his pronouncements against those of others. He does not have the critical conceptual resources available to distinguish the *pragmatics* of subjectively resonant philosophy and propaganda. After acknowledging the need to make a distinction between genuine and counterfeit contact with moral sources, and the dangers of forgetting it, he offers nothing by of justifying such a distinction.

The objective critic balks at the subjective critic's quietism on this issue. More specifically, Habermas can argue for the need to establish intersubjective constraints on the project of world-disclosure. The argument would look something like this. The pronouncements of the poet exploit the world-disclosive powers which are *intrinsic* to language in its capacity to mediate rationally coordinated action, even though the capacity for world-disclosure is enhanced by the bracketing of illocutionary obligations. The speech-act pragmatics of propaganda, on the other hand, exploit extra-linguistic coercive forces for the purpose of persuading the hearer of the acceptability of what is disclosed in the propaganda. Of course, Habermas elaborates an argument of this kind in his construction of the conditions of the intelligibility of systematically distorted communication, and thereby of the intelligibility of a critique of ideology. This kind of critique thus falls within the province of objective criticism. We can agree with Habermas that the non-coercive pragmatics of epiphanic poetry or philosophy implicitly presuppose the symmetrical intersubjective structure of communicative action. But this does not commit us to the claim that the legitimate task of the subjective critic is limited to the raising and redeeming of validity

claims to authenticity. Nor does it commit us to the assertion that the validity of subjective criticism is criterially constrained by the possibility of reaching a consensus. As I have argued, Habermas's taxonomy of validity claims, as well as his views on the consensual orientation of communicative action, must be rejected. But these claims are adventitiously related to the model of democratic will-formation represented in the dialogical situation of communicative problem-solvers; a situation which it is not unreasonable to suppose can structure the public order of references under conditions of modernity.

Finally, I want to stress that the tasks of objective and subjective criticism as I am introducing them here are neither 'objectivist' nor 'subjectivist', insofar as these terms have accrued pejorative connotations. The reason for this becomes manifest by considering their contribution to defining the critical thrust of political ecology. Subjective criticism as perpetrated and advocated by Taylor has a realist core; it is *about* subjectivity only in the sense that the world is disclosed only under certain conditions of intentionality, but in articulating these conditions, its critique is directed towards the world. Subjectivity is understood as the medium of normative orientations, not their (exclusive) source; it is subjective in "*manner*", not in "*matter*".¹³ To say that the natural environment, as a world disclosed in languages with various potentials for personal resonance, must matter *to* human beings, is not necessarily to claim that human beings are the only things that matter. The pejorative force of 'subjectivism' is attendant first upon the *ontological* claim that value exists in the world just to that extent to which a subject's de facto desires are satisfied, and following from this, upon the *ethical claim* that the best world is one in which all of these (compossible) desires are in fact satisfied. Since subjective criticism is neither ontologically nor evaluatively subjectivist, it is not a fit object of the pejorative force associated with the

term 'subjectivism'. It is, however, semantically subjectivist, but only with respect to a restricted domain of discourse. This is the domain in which meaning has an irreducibly experiential core, which in turn is an apt object of phenomenological exploration. This is the domain of 'private' language only to the extent that, as a matter of fact, the claims expressed in it are not recognized in what Taylor calls the public order of references. The language of subjective criticism is not private in the empiricist sense - attacked by Wittgenstein - of being meaningful in virtue of referring to a private ontology of internal mental states. I see no reason why this limited semantic-phenomenological subjectivism should offend the anti-anthropocentric sensibilities of the radical ecologist.¹⁴

Just as the innocence of subjective criticism from the charge of 'subjectivism' is testified by its link with the politics of world-disclosure, so the connection between objective criticism and the politics of democratic problem-solving can disarm worries about the complicity of such critics with 'objectivism'. The objectivity secured by intersubjective agreement under communicative constraint concerns a problem-solving capacity for rational action co-ordination. Again, the pejorative force of 'objectivism' is directed towards certain ontological, epistemological, and ethical commitments: that the world is nothing but an aggregate of atomic objects interacting according to mechanistic laws; that knowledge of it is exclusively a function of a method of the kind which modern physicists are supposed (by objectivist philosophers) to follow; and that it is good to follow it - both as an expression of human excellence in proper intellectual conduct, and in its utility. But if we think of the objective critic, along Habermasian lines, as clarifying the public or intersubjective conditions of rational action co-ordination, the objectivist charge has no real force. Indeed, objective criticism comes into play when the intersubjectively binding

presuppositions of communicatively rational action co-ordination are violated by the pseudo-objectifications of science.¹⁵

How, then, can it come into play in the kind of critique which is relevant to the radical ecologist? One suggestion which has been made goes like this; the critical thrust of communicative reason can be salvaged from its anthropocentric bias by including the rich diversity of life-forms - and even the earth itself - as virtual *participants* in the dialogue situation.¹⁶ But it is extremely difficult to see how this idea could at all be sustained; and impossible to see how it could fit with anything remotely resembling communicative ethics.¹⁷ A much more plausible idea is hinted at, but not followed through, by Habermas himself. As I said, objective criticism comes into play when the intersubjectively binding presuppositions of communicative action are violated. But communicative action presupposes the integrity of a background lifeworld horizon. Further, as Habermas remarks above, the integrity of the *organic foundations* of the lifeworld can come, indeed are, under threat. The violation of these organic foundations, it follows, is a suitable object for objective criticism; their breakdown is criticizable (objectively) in virtue of being inconsistent with the presuppositions of unconstrained communicative action as a medium of will-formation. This doesn't tell us *why* the environment itself might be intrinsically worth preserving, independent of what it makes possible for humans. But then to show that is the world-disclosive task of the subjective critic.

The critical thrust of political ecology can thus be understood as standing at the interface of the politics of world-disclosure and problem-solving. This point can be seen to correspond to, but is not identical with, Habermas's understanding of green politics as a

border conflict between system and lifeworld. According to Habermas, green protest is a reaction from the point of view of the lifeworld to an environmental crisis which manifests itself as a visible attack upon the organic foundations of the lifeworld. But it also generates "a new category of literally invisible risks that can only be grasped from the vantage point of the system".¹⁸ These are risks the responsibility for which is offset by their 'uncontrollable magnitude'. They are problems which must be tackled with the resources of the lifeworld, yet, as Habermas remarks, they go beyond the level of complexity which can be experienced within it. So to the *liberating* riskiness of self-transfiguration which the subjective critic celebrates, is now added the *constraining* risk of responsibility for action-consequences of hitherto incomprehensible moral magnitude.

While the politics of world-disclosure struggles to expand and deepen the resources of the lifeworld, the politics of democratic problem-solving pursues means of putting a heightened sense of responsibility into effect by way of a particular form of decision-making.¹⁹ The success of the critique of instrumentalism, and the prospects for modernity's self-reassurance, turns on the headway made in both these directions. Through the creation of new languages of personal resonance, it can disclose matterings in the world, or voice previously inarticulate moral claims, which must be capable of *motivating collectively co-ordinated action* constrained by the social and ecological imperative of effective democratic problem-solving.

NOTES

Chapter One: Introduction.

1. In Karl Marx, 'Letter to A. Ruge', September 1843. Cited in Nancy Fraser, 'What's Critical about Critical Theory? The Case of Habermas and Gender', *New German Critique*, 35, Spring/Summer 1985, p97.
2. These questions might be put by someone - perhaps sympathetic to the 'analytical' tradition of philosophy - who holds that since 'modernity' is an empirical phenomenon, it is something to be investigated exclusively by means of the empirical method of the social sciences. Such enquiry, on this view, would be philosophically competent only to the extent that it avoided conceptual confusion, and certainly would not be philosophical in content. On the other hand, the question might be put by someone - more in touch with the 'Continental' tradition - who holds that philosophy has in some sense 'come to an end', superseded, for instance, by revolutionary praxis, meditative thinking, or deconstruction. While superficially both these views can be seen as simply opposing one definition of philosophy to another, at a deeper level the tension between them can be understood as part of the very 'modernity' in question; for as conceptions of rationality change with modernity, so do conceptions of philosophy. For a highly accessible overview of the various 'ends' of philosophy in historical context, see Harry Redner, *The Ends of Philosophy: An Essay in the Sociology of Philosophy and Rationality*.
3. Exemplary works in this area are: Martin Jay, *Marxism and Totality: Adventures of a concept from Lukacs to Habermas*; Seyla Benhabib, *Critique, Norm, and Utopia: A Study of the Foundations of Critical Theory*; and Axel Honneth, *The Critique of Power: Reflexive Stages in a Critical Social Theory*.
4. These studies are now many, but particularly noteworthy are: Jürgen Habermas, *The Philosophical Discourse of Modernity; Twelve Lectures*, lectures I-IV; Charles Taylor, *Hegel*; David Kolb, *A Critique of Pure Modernity*; and Robert Pippin, *Modernism as a Philosophical Problem*.
5. This focus is at its clearest in Peter Dews, *Logics of Disintegration; Post-structuralist Thought and the Claims of Critical Theory*, and in the Foucault chapters of Honneth's *The Critique of Power*.

6. Habermas's response is developed in *The Philosophical Discourse of Modernity: Twelve Lectures*, lectures VII-XII. For rejoinders, see: Jay Bernsetin, 'The Causality of Fate: Modernity and Modernism in Habermas', *Praxis International*, 8, January 1989 (special issue devoted to Habermas's book) pp407-25; John McCumber, 'Philosophy as the Heteronomous Center of Modern Discourse: Jürgen Habermas', in H.Silverman (ed.) *Philosophy and Non-Philosophy Since Merleau-Ponty*, pp211-231; Tony Cascardi, *The Subject of Modernity*; and less antipathetic to Habermas, Martin Jay, 'The Debate over Performative Contradiction: Habermas vs. the Post-structuralists', in T.McCarthy et.al. (eds.) *Zwischenbetrachtungen im Prozeß der Aufklärung*, pp171-189. For wide-ranging discussions aimed at bridging the gap between Habermas and the post-structuralists, see: Thomas McCarthy, *Ideals and Illusions: On Reconstruction and Deconstruction in Contemporary Critical Theory*; Richard Bernstein, *The New Constellation: The Ethical-Political Horizons of Modernity/Postmodernity*; and Stephen K. White, *Political Theory and Postmodernism*. And for general introductory overviews, see: Martin J. Matustik, 'Jürgen Habermas at 60', *Philosophy and Social Criticism*, 16, January 1990, pp61-79; David Rasmussen, *Reading Habermas*, ch.6; and Robert C. Holub, *Jürgen Habermas: Critic in the Public Sphere*, ch.6.
7. Both Taylor and MacIntyre take Foucault to be the most challenging representative of post-structuralism; see Taylor, 'Foucault on Freedom and Truth' in his *Philosophy and the Human Sciences*, pp152-84, and MacIntyre, *Three Rival Versions of Moral Enquiry*, especially chapters II and IX (containing just one critical reference to Habermas, on p46). As yet, Taylor's only direct encounters with Habermas's work are: 'Language and Society', in Honneth and Joas (eds.) *Communicative Action*, pp23-35; *Sources of the Self*, pp85-88 and 509-10; and 'Comments and Replies', *Inquiry*, 34, June 1991, pp250-3. Habermas has responded briefly in 'A Reply', published in the same volume as Taylor's article, pp215-22. He has written a more extensive response to the kind of position Taylor defends in two articles; 'On the Pragmatic, the ethical and the moral employments of practical reason', and 'Lawrence Kohlberg and Neo-Aristotelianism', translated by Ciaran Cronin, to be collected in *Remarks on Discourse Ethics* (Polity; Cambridge) forthcoming.
8. For comparisons of MacIntyre and Habermas, see: Michael Kelly, 'MacIntyre, Habermas and Philosophical Ethics', in *The Philosophical Forum*, XII, Fall-Winter 1989-90, pp70-93, and William Rehg, 'Discourse Ethics and the Communitarian Critique of Neo-Kantianism', in the same journal, XXII, Winter 1990, pp120-138. The latter article contributes to a debate between so-called 'communitarians' and 'liberals' or 'universalists'. It is in the

context of this debate that the ('liberal/universalist') claims of Habermas and the ('communitarian') claims of Taylor have been compared. For a representative sample of the issues at stake, see David Rasmussen (ed.), *Universalism vs. Communitarianism*. While these issues are important, the playing-off of 'liberal' and 'communitarian' labels has lost from view the contrasting philosophical conceptions of modernity, crisis and critique which underlie both Habermas's and Taylor's views on morality and politics. Ross Poole effects such a contextualization of MacIntyre's views in *Morality and Modernity*, though without much detail on Habermas, and with no reference to Taylor.

James Tully, in 'Wittgenstein and Political Philosophy; Understanding Practices of Critical Reflection', *Political Theory*, 17, 1989, pp177-204, criticizes both Habermas and Taylor on Wittgensteinian grounds - as I do later in the thesis. Paul Redding's 'Habermas's Theory of Argumentation', *The Journal of Value Inquiry*, 23, 1989, pp15-32, criticizes Habermas by reference to ideas drawn partly from Taylor's work, reaching conclusions with which, as will become clear, I am in sympathy.

9. Taylor, 'Inwardness and the Culture of Modernity', in T.McCarthy et.al. (eds.) *Zwischenbetrachtungen im Prozeß der Aufklärung*, pp601-23.
10. Ibid., p601.
11. MacIntyre, 'Epistemological Crises, Dramatic Narrative, and the Philosophy of Science', *The Monist*, 60, 1977, p453. In his 'Explanation and Practical Reason', *Wider Working Papers*, 1989, Taylor draws heavily upon the account given by MacIntyre in this article for his own conception of practical reasoning. I discuss it in chapter seven.
12. This canon being, more or less, that systematically reconstructed in Habermas's *The Theory of Communicative Action*; Weber, Durkheim, Mead, Parsons, Marx and Lukács.
13. Nietzsche, for instance, has often been criticized for exaggerating the value accorded to pity in the Christian ethic, and a similar charge has been put against Foucault's (over)emphasis on the 'disciplinary' contortions of the modern self. The general objection is put by Taylor in *Sources of the Self*, p503. By focussing exclusively on just one ideal, these approaches tend to fall into the error of thinking that the undesirable consequences of the exclusive pursuit of any single ideal *invalidate* that ideal, Taylor argues.
14. Taylor, 'Inwardness and the Culture of Modernity', p602.
15. Of course, the same theorist can hold both these views.
16. See Taylor, 'Neutrality in Political Science', *Philosophy and the Human Sciences*, pp58-90 for an early elaboration of this claim.
17. Taylor, 'Inwardness and the Culture of Modernity', p605.

18. Being Habermas's theorization of modernity as differentiated between the 'value spheres' science, morality, and art, mentioned below.
19. See Habermas, *The Theory of Communicative Action: Vol.1*, pp6-7.
20. Habermas, *The Philosophical Discourse of Modernity*, p4. In the summary which follows, I draw on the what Habermas self-effacingly calls 'thought experiments' in *ibid.*, pp342-48. These are very condensed summaries of the argument developed throughout both volumes of *The Theory of Communicative Action*, but especially in Vol.2, p140f.
21. Here I draw on 'Modernity - An Incompleted Project', and *The Philosophical Discourse of Modernity*, p340.
22. 'Modernity - An Incompleted Project', p9.
23. This is a claim which undergoes several different modifications in the Habermas opus. Most important, in his more recent work (eg. *The Philosophical Discourse of Modernity*, p389) he has wanted to stress the 'world-disclosive' capacity of art. But it is difficult to see how this capacity can be grasped under the rubric of rationalization. How it might be understood is an issue I take up in chapter eight.
24. Habermas, *The Philosophical Discourse of Modernity*, p16.
25. Habermas, *The Theory of Communicative Action, vol.1*, p3
26. Habermas, *The Philosophical Discourse of Modernity*, p326.
27. See Habermas, *Legitimation Crisis*, pp1-2.
28. *Ibid.*, p1.
29. *Ibid.*
30. *Ibid.*
31. *Ibid.*, p2.
32. My thesis that there are 'two Habermases' competing with each other thus converges with the claim made by Axel Honneth in *The Critique of Power*, that Habermas shifts between two different models of disturbed will-formation in the evolution of the human species; one provoked by the 'technocracy thesis' - that the human species develops its rational capacities solely to the extent to which it finds technical solutions to problems facing it - and one based on 'the dialectic of moral life' which assumes the form of a struggle for recognition between social groups or classes (pp268-77). While I am broadly sympathetic to Honneth's interpretation, and to his preference for the second model, he doesn't consider the philosophical implications, central to my argument, of the link between crisis as conceptualized in the second model, and the form of critique appropriate to it. See my review essay of Honneth's book in *History of the Human Sciences*, forthcoming.

33. Habermas, *The Philosophical Discourse of Modernity*, p16
34. Ibid., p26.
35. Ibid., p20.
36. Ibid., p139.
37. The distinction between 'systemic' and 'lived' crises is elaborated by Benhabib, in *Critique, Norms and Utopia*, pp123-33.
38. The acknowledgement to Gadamer comes in Taylor, 'Understanding and Explanation in the *Geisteswissenschaften*', p206
39. Gadamer's main contributions to the debate are: 'The Universality of the Hermeneutical Problem', in Joseph Bleicher (ed.) *Contemporary Hermeneutics* (with commentary by the editor) and 'The Scope and Function of Hermeneutic Reflection', in Gadamer, *Philosophical Hermeneutics*, pp18-43. Important presentations of Habermas's counter-claims include: 'The Hermeneutic Claim to Universality' (also in Bleicher); two articles in *Inquiry*, 13, 1970, 'On Systematically Distorted Communication', and 'Towards a Theory of Communicative Competence'; and the essay 'Interpretive Social Science vs. Hermeneuticism' in N.Haan et.al. (eds.) *Social Science as Moral Enquiry*. Extensive secondary literature occasioned by the debate is thoroughly documented in Robert C. Holub, *Jürgen Habermas: Critic in the Public Sphere*, pp193-4
40. See Gadamer, *Truth and Method*, p238f., and 'The Universality of the Hermeneutical Problem', p9.
41. *Truth and Method*, p273f.
42. See Richard Bernstein, *Beyond Objectivism and Relativism*, p8f. for an illuminating and influential discussion of its significance.
43. Habermas, 'The Hermeneutic Claim to Universality', p203.

Chapter Two: The Modern Tension between Cognition and Identity

1. Gellner, *Thought and Change*, p 72. This new balance is supposed to be congruent with the industrialization of society, where "industry is, essentially, the ecology of a society possessed by science", and "science, essentially, is the form of cognition of industrial society". Fred Dallmayr complains of Habermas's failure in *The Philosophical Discourse of Modernity* to comment on the contribution of the Anglo-Saxon tradition of philosophy to the entitled discourse ("The Discourse of Modernity: Hegel, Nietzsche, Heidegger (and

Habermas)', *Praxis International*, 8, Jan. 1989, p397). I discuss Gellner at this point not least because *Legitimation of Belief* can fruitfully be read as filling this gap. Empiricism, logical positivism, and ordinary language philosophy are presented as more or less successful attempts at providing self-reassurance for the modern identity, as expressed in what will be called the ethic of cognition.

2. Gellner, *Legitimation of Belief*, p207.
3. Ibid., pp27-9, *Spectacles and Predicaments*, p164f.
4. Gellner, *Legitimation of Belief*, p27. Gellner cites the following passage from the Renaissance philosopher Giordano Bruno as an early 'specimen' of this thought;

Saul: And let me hear why Jove has decided that Truth shall be placed in the highest seat.

Sophia: Easily. Truth is placed above all things...for if we were to conceive of something that was to be ranked higher, it would also have to be something other than Truth. And if you imagine it to be something other than Truth, you will necessarily see it as not being true, hence it is false, it is worse than nothing.

This sense of the indignity of holding false beliefs, the fear of shame at being duped, helps to explain the virtue which is accorded to holding beliefs only under maximal epistemic constraints.

5. Ibid., p207.
6. Ibid., p174.
7. Ibid., p180.
8. Ibid., p37. Gellner's comment refers specifically to Hume's prescription for consigning volumes of metaphysics to the flames, but it is equally applicable to Locke's epistemological project.
9. A thesis which converges closely with Rorty's in *Philosophy and the Mirror of Nature*, pp131-9. Here, as on many other points concerning the significance of epistemology, Gellner anticipates Rorty. The divergence lies in the moral of their stories; Gellner wants to praise epistemology, Rorty to bury it.
10. Gellner, *Legitimation of Belief*, p100.
11. Ibid., p63.
12. Gellner, *Culture, Identity, and Politics*, p153.
13. Gellner, *Legitimation of Belief*, p100.
14. See McDowell, *The John Locke Lectures*, 1991, unpublished manuscript, pp56-7.

15. Ibid., p60. McDowell calls this view 'bald naturalism', to which he opposes two alternative positions. First, one based on Davidson's 'supervenience' thesis, and second - his own preferred view - a 'naturalized platonism', according to which concepts suitable to the realm of meaning (rather than the disenchanted realm of law) constitute an 'openness to the world'. McDowell draws on the work of Gadamer in order to elaborate and defend this notion of openness to the world, and though I cannot deal with his arguments here, suffice it to note that his conclusions are very close to those urged by Taylor, whose arguments will be examined.
16. Gellner, *Legitimation of Belief*, p106.
17. Ibid. p196.
18. Ibid.
19. Gellner, *Spectacles and Predicaments*, p60.
20. Ibid.
21. Gellner, *Relativism and the Social Sciences*, p77f.
22. Ibid., pp81-2.
23. Husserl, *Crisis in the European Sciences*, p142. I would like to make a few remarks about what the 'life' of the lifeworld connotes. First, there is the connotation of a domain which cannot be fully grasped by purely theoretical reason. "Only life understands life", Dilthey remarks. (Quoted in Garbis Kortian, *Metacritique*, p129). In addition to denoting the pre-theoretical and the pre-reflective, there is an appeal to the unfathomability of human existence as a whole. The connotation is that of the irreducible complexity, contingency, and significance of an individual or collective life which can only be grasped, if at all, "from within" by those who live it. Second, there is a normative connotation to be drawn from 'life'. Here, life connotes vitality. In the next chapter we will see that this is explicit in the young Hegel's account of an ethical totality subject to a 'causality of fate'. For an individual or collective to lose the vitality of 'life' is for it to lose something distinctly human, to fall short of an ethical standard of a fully realized humanity. Third, life connotes meaning. This connotation is suggested by Wittgenstein in his remark that a sign in itself, as merely a written mark or vocalization, appears as "utterly dead" (*The Blue Book*, p4). Seeing things this way, Wittgenstein suggests, can tempt one into looking for "that which gives the sentence life as something in an occult sphere" which co-exists with the signs. (*The Blue Book*, p5). We can avoid this picture, he suggests, if we take the life (i.e. the meaning) of the sign to lie in its "use". But even if in its use the sign is alive, Wittgenstein asks in *Philosophical Investigations*, "Is life breathed into it there? - Or is the use its life?" (#432)

Those who take an expressivist theory of meaning from Wittgenstein favour the former option, those who take a use theory from him the latter. However interpreted, Wittgenstein captures a non-trivial element in our intuitive sense of meaning by indicating how we link it to the 'life' of a sign.

The second and the third connotation are, of course, connected. Reference to a 'dead language' betokens a diseased and decadent culture. The moral imperative of preserving meaning in language is perhaps put most radically by Karl Krauss, in a remark approvingly cited by Anton Webern; "[moral gain] lies in a spiritual discipline which ensures the utmost responsibility toward the only thing there is no penalty for injuring - language" (*The Path to the New Music*, pp9-10). Avant-garde art is also unthinkable without such a commitment, given powerful articulation in Octavio Paz's comment that "When a society becomes corrupt, what first grows gangrenous is language. Social criticism, therefore, begins with grammar and the re-establishment of meanings" (*On Poets and Others*, pxiii). Both Habermas's and Taylor's philosophical conceptions of crisis and critique fall very much within this modernist tradition.

24. Husserl, *Crisis in the European Sciences*, p130.
25. As the 'given', the lifeworld was to feature as the point of departure for a resurrected transcendental phenomenology. But as Husserl sensed, it is difficult to see how the phenomenological method of intuitive reflection aiming at absolute a priori apodeictic truth can be reconciled with the intersubjective insights captured in an essentially historicist lifeworld concept. This methodological difficulty can be overcome, claims Habermas, only by a thoroughgoing shift to intersubjectivity in which the subject-object figure of thought itself is made derivative of the prior intersubjectivity of a linguistified lifeworld, and 'reductions' of consciousness are replaced by structures of linguistic communicative competence.
26. Habermas, *The Philosophical Discourse of Modernity*, p147.
27. Habermas, *Theory of Communicative Action*, vol.2, p124.
28. Wittgenstein, *On Certainty*, #162.
29. *Ibid.*, #94.
30. Habermas, *Theory of Communicative Action*, vol.1, p70.
31. Habermas, *Theory of Communicative Action*, vol.2, p130.
32. *Ibid.*, p126.
33. Gellner, *Spectacles and Predicaments*, introduction.
34. Habermas, *The Theory of Communicative Action*; vol.1, p100; vol.2, p120.

35. The validity claims of truth, rightness, and authenticity are thematized in the discourses of the value spheres of science and technology, law and morality, and art and criticism; this differentiation and reflexive thematization partly constitute cultural rationalization. The lifeworld shares an analogous three-fold formal structure, which Habermas calls 'culture', 'society', and 'personality'.
36. Habermas, *The Theory of Communicative Action*, vol.1, p340.
37. Ibid., p10.
38. Habermas, *The Philosophical Discourse of Modernity*, p311.
39. Ibid., p315-6. I return to this theme in chapter eight.
40. "There is a reason why the world the world is made of machines, and that reason lies not in the world but in our practices of explanation", *Legitimation of Belief*, p.107.
41. Wittgenstein, *Remarks on the Foundations of Mathematics*, I. 93.
42. Ibid., IV. 48.
43. Since taken this way; the rule admits of an infinite number of possible interpretations; the conclusion reached by #201 of *Philosophical Investigations*.
44. The epistemological significance of primary qualities is, of course, that in being measurable and quantifiable, the verification procedure for truth-claims in which they feature is minimally vulnerable to sceptical objection.
45. See Taylor, 'Understanding in the Human Sciences', *Review of Metaphysics*, 1980, 34, pp25-38; and the subsequent exchange with Rorty (pp47-55).
46. See John McDowell, 'Virtue and Reason', *The Monist*, 62, 1979, pp331-350 (especially p345f.).
47. This is the first premiss of an argument which could easily lead to the conclusion that we are referring here to an eroding power which is not intellectual, but ideological.
48. See Taylor, *Sources of the Self*, p.8.
49. As is the case, for instance, when anthropologists are confronted with apparently irrational primitive societies. The problem is discussed by Taylor in his paper 'Rationality', *Philosophy and the Human Sciences*, pp134-51.
50. *Legitimation Crisis*, p120
51. At least in one of his phases; see Dews, *Logics of Disintegration*, pp203-6.
52. Nietzsche, 'On Truth and Falsity in an Ultramoral Sense', in *The Philosophy of Nietzsche* (ed.) Geoffrey Clive, p508.

53. Habermas quotes Nietzsche remarking in *The Birth of Tragedy* of the "terror which seizes man when he suddenly begins to doubt the cognitive form of phenomena, because the principle of sufficient reason ..seems to suffer an exception. If we add to this terror the blissful ecstasy that wells from the innermost nature of man, indeed of nature, at this collapse of the *principii individuationis*, we steal a glimpse into the nature of the Dionysian" (Habermas, *The Philosophical Discourse of Modernity*, p93). According to this reading of Nietzsche, redemption is possible for an impoverished mythless modernity in the form of aesthetic experiences of Dionysian ecstatic self-oblivion. Countering this, Habermas suggests that this kind of appeal to the redemptive power of art will not do because the aesthetic phenomena of which Nietzsche speaks, far from being of pre-modern mythic origin, are in fact only possible given the differentiation of art from cognition and morality which is the achievement of modernity. They are the kind of experiences explored by artistic modernism, and more specifically, the avant-garde. Habermas's emphasis is slightly different to the one I am making here. But I will return to the aesthetic potential for the reconciliation between the diremptions of modern life in both Habermas and Taylor in chapter eight.
54. The way this is interpreted is decisive for the different approaches to the philosophical problem of modernity. We will see that even Taylor agrees with the point, but *what* reality manifests itself through the aesthetic creation of humans is treated very differently.
55. Nietzsche, *The Will To Power* translated by W. Kaufmann, #552. The idea of truth as invention, and the ethic of self-creation which "in itself has no end", but without the will to power metaphysics, define exactly the project of Richard Rorty's recent writings. The basic thought of the norm of the endless redescription of ourselves which Rorty advocates is itself but a redescription of this Nietzschean theme.
56. Habermas, *The Philosophical Discourse of Modernity*, p296.
57. Taylor, *Sources of the Self*, pp98-103.

Chapter Three: Habermas's Conception of Crisis

1. Habermas, *The Philosophical Discourse of Modernity*, p348.
2. Ibid., p316.
3. Ibid., p306.
4. See Habermas, *Theory and Practice*, pp28-9.

5. In Hegel, *Early Theological Writings*, pp182-301. As will become clear, Hegel's critique of the idea that what might be called "the pull of the ethical" is that of obligation to the moral law is largely directed against Kant's views.
6. Ibid., p226.
7. Ibid., p238.
8. Ibid., p231. The issue is complicated by the role of Kant's 'holy will' which transcends such inclinations. Kant seems to be left in the odd position that from the point of view of the moral worth of the will, it is actually good fortune to have such potentially immoral natural inclinations, since it gives the will all the more opportunity to prove its worth. Hegel's underlying point is that the avenging force of the ethical stands outwith the autonomy of the isolated rational will - and to the degree to which it is therefore heteronomous, the Kantian opposition between the autonomous and the heteronomous breaks down.
9. Ibid., p232. This reference holds for the remaining quotations in this paragraph.
10. Ibid., p228.
11. In Williams, *Ethics and the Limits of Philosophy*, pp174-196.
12. Ibid., p193.
13. This is slightly misleading. The universality of the moral law has its foundation in the necessity of practical reasoning, as Williams explains in chapter four of his book.
14. Ibid., p191 [my italics].
15. Ibid., p195.
16. Ibid., p193.
17. Ibid., p196.
18. Williams gives us reason for thinking that these oppositions need to be reconciled, but offers little by way of indicating how they might be. Elsewhere, he acknowledges that an answer might lie in a tragic conception of agency (*Moral Luck*, p30, footnote 2). By conceiving the force of the ethical as fate rather than as luck, Hegel is able to develop such a conception systematically.
19. Hegel, *Early Theological Writings*, pp230-1. The feeling of a life disrupted 'must' become a longing for what has been lost if there is to be reconciliation, but the force of this 'must' is not categorical; it is not determined by the moral law.
20. Ibid., p231.
21. Ibid., p233.

22. As in tragedy, even the innocent gets caught in the nexus of guilt representing the avenging force of life (the friendliness of collective life together).
23. Ibid., p238.
24. Ibid., p233.
25. Ibid., p230.
26. Ibid., p229.
27. Ibid.
28. Ibid., p232. Despite his hostility to the christian ethic, even Nietzsche is sometimes tempted to take a similar position; "That which is done out of love always takes place beyond good and evil". (*Beyond Good and Evil* #153). The problem is that Nietzsche is so obsessed with what he sees as christianity's 'slave morality' based on 'revenge' or '*ressentiment*' to recognize Hegel's point. In the *Genealogy of Morals*, Nietzsche even seems to suggest that it is precisely lack of recognition of other parties which constitutes *ressentiment*; "the slave morality" says "no" from the very outset to what is "outside itself", and "not itself"; and this "no" is its creative deed. This *volte face* of the valuing standpoint - this inevitable gravitation to the objective instead of back to the subjective - is typical of "*resentment*". This remark is a very nice illustration of the critique of what Habermas calls 'subject-centered reason' which unwittingly reproduces that kind of reason. Unlike Nietzsche, Hegel avoids the oscillation between the objective and the subjective, since subjects and objects only emerge out of a prior context of intersubjectivity. And it is the revenge of the distortion of *this* shared context, so Hegel is arguing, which is the true 'genealogy' of morals, rather than the revenge of a self-defensive or 'reactive' objectifying 'herd'.
29. First in 'Labour and Interaction: Remarks on Hegel's Jena Philosophy of Mind', in *Theory and Practice*, pp142-169, especially pp150-6. He later defends his own quasi-Kantian 'discourse ethics' as innocent of these weaknesses in 'Morality and Ethical Life: Does Hegel's Critique of Kant apply to Discourse Ethics?', in Habermas, *Moral Consciousness and Communicative Action*, pp 195-215, a text which I discuss in chapter five.
30. Habermas, *Theory and Practice*, p148. The same passage is included in *Knowledge and Human Interests*, p56 (under a different translation).
31. Ibid.
32. Habermas, *The Philosophical Discourse of Modernity*, p324.
33. Honneth discusses how, in the early chapter of *Knowledge and Human Interests*, Habermas incorporates the causality of fate theme within a Marxian framework of the self-formative

process of the species in *The Critique of Power*, pp268-77. But as I go on to indicate, the *kind of theory* (namely *Critical Theory*) which Habermas thinks can best articulate it is most thoroughly discussed in the later chapters on Freud.

34. Ibid., p316.
35. Especially chapters 10 and 11. My exposition also draws on Habermas's 'The Hermeneutic Claim to Universality', in Bleicher (ed.) *Contemporary Hermeneutics*, pp181-211.
36. Habermas, *Knowledge and Human Interests*, p229.
37. And of course, words can be dangerous. One may not trust one's partner in communication, one may not even trust *oneself* in communicating.
38. "To imagine a language means to imagine a form of life". (Wittgenstein, *Philosophical Investigations*, #19). I am of course only asserting the conclusions here of Wittgenstein's arguments, but I will be returning to them in chapter four.
39. Habermas, *Knowledge and Human Interests*, p226.
40. Ibid., p227.
41. The difficulties of this thought are taken up by Donald Davidson, 'Paradoxes of Irrationality', in Woolheim and Hopkins (eds.), *Philosophical Essays on Freud*, pp289-305. The very different lesson for moral reflection which Rorty gets from Freud through Davidson's paper will be discussed in (6.4) below.
42. Habermas, *Knowledge and Human Interests*, pp216-7.
43. Habermas, 'The Hermeneutic Claim to Universality', p191.
44. Ibid., p194.
45. Habermas, *Knowledge and Human Interests*, p228.
46. Ibid., p256.
47. Ibid., p241.
48. Ibid., p256.
49. Ibid., p271.
50. Ibid.
51. Ibid., p245.
52. Ibid., p257.
53. Habermas, 'The Hermeneutic Claim to Universality', p193.
54. Habermas, *Knowledge and Human Interests*, p232.
55. Ibid., p344-5 footnote 35.
56. Ibid., p234.
57. Ibid., pp235-6.

58. Especially chapters 4,5,6. Keat's book was published in 1981, since when Habermas has modified many of the positions attacked by Keat.
59. Keat, *The Politics of Social Theory*, p108.
60. Ibid., p107.
61. Habermas, *Knowledge and Human Interests*, p271. cited in Keat, p.107.
62. Keat, *The Politics of Social Policy*, p108.
63. Habermas, *Knowledge and Human Interests*, p220.
64. Habermas acknowledges that for Freud, this is an inevitable part of human psycho-social development.
65. Keat, *The Politics of Social Theory*, p179. After making this comment, Keat turns to Habermas's attempt to clarify the rational foundation of norms in discourse. Keat fails to appreciate that the model of communicative action through which Habermas attempts to do this is *already* central to his interpretation of pschoanalysis.
66. Habermas, *Knowledge and Human Interests*, p232.
67. J.M.Bernstein, 'Self-Knowledge as Praxis: Narrative and Self-Narration in Psychoanalysis', in Christopher Nash (ed.) *Narrative in Culture*, p60.
68. Ibid., p62.
69. For this reason, Bernstein can put the objection to Habermas that he perpetuates the "Enlightenment" illusion that "*who the self is and its freedom can be analytically distinguished*" (ibid., p75).

Chapter Four: Habermas, Language, and Critique

1. Habermas, 'Reply to my Critics', in John B.Thompson and David Held (eds.), *Habermas: Critical Debates*, p227.
2. Ibid.
3. As Habermas puts it in *The Philosophical Discourse of Modernity*, p315:
both cognitive-instrumental mastery of an objectivated nature (and society) and narcissistically overinfalted autonomy (in the sense of purposively rational self-assertion) are derivative moments that have been rendered independent from the communicative structures of the lifeworld, that is, from the intersubjectivity of relationships of mutual understanding and relationships of reciprocal recognition. Subject-centred reason is the product of division and usurpation, indeed of a social process in the course of which a subordinated moment assumes the place of the whole, without having the power to asimilate the structure of the whole".

4. See for example John B. Thompson, *Critical Hermeneutics*, pp130-9; Fred Dallmayr, 'Life-World and Communicative Action', in *Polis and Praxis*, pp224-54; and Nancy Fraser, 'What's Critical about Critical Theory? The Case of Habermas and Gender', *New German Critique*, 35, Spring/Summer 1985, pp97-131 (especially pp100-5)
5. As Habermas implies in *The Theory of Communicative Action, vol.1*, p397.

If we assume that the human species maintains itself through socially coordinated activities of its members and that this coordination has to be established through communication and in central spheres through communication aimed at reaching an agreement - then the reproduction of the species also requires satisfying the conditions of a rationality that is inherent in communicative action.
6. *Ibid.*, p286.
7. More importantly, Habermas proposes that this de-differentiation occurs in the form of systematically distorted communication. "Defense mechanism", Habermas writes, "undermine the differentiation between action oriented to success and action oriented to reaching understanding", by "separating the (unconscious) strategic aspect of action (which serves the gratification of unconscious desires) from the manifest intentional action that aims at reaching understanding." *Moral Consciousness and Communicative Action*, p188.
8. *Ibid.*, p133. On the objection that what we would normally call moral (non ego-utility maximizing) considerations may also feature in the strategy of the actor, see Stephen K. White, *The Recent Work of Jürgen Habermas*, pp10-11.
9. *The Theory of Communicative Action, vol.1*, p286.
10. Arie Brand illustrates the distinction between strategic and communicative action with an imaginative example of this kind in his *The Force of Reason*, pp14-17. I take up the objection that these examples of strategic action are distortingly one-sided below.
11. *The Theory of Communicative Action, vol.1*, p301.
12. *Ibid.*, p288.
13. And as observed in Note 7 above, the unconscious can be understood as playing the role of a 'strategic actor occupying "internal foreign territory"'.
14. Stephen K. White, *The Recent Work of Jürgen Habermas*, p46. I am proposing that at the *level of justification*, the category of normativity by which social life can be criticized, and which gives content to the idea of *undamaged* social life within Habermas's theoretical framework, requires that priority be given to one of these two kinds of *linguistic* interaction.

15. Ibid. To uphold a distinction between 'original' and 'parasitic' or 'derivative' modes of language use is not necessarily to claim that communicative action is more fundamental to social life than *purposive* action. Since composing *The Theory of Communicative Action*, Habermas has abandoned the latter position (if he ever held it) while retaining the former. And it is this former claim about *linguistic* interaction which, I am proposing, against White's interpretation, remains crucial. See Habermas's 'A Reply' in Honneth and Joas (eds.) *Communicative Action*, p237, and Habermas, *Postmetaphysical Thinking*, pp82-4, and pp86-7; note 33.
16. Rasmussen, *Reading Habermas*, ch.2.
17. See Jonathan Culler, 'Communicative Competence and Normative Force', and Allan Wood 'Habermas' Defence of Rationalism', both in *New German Critique*, 35, 1985.
18. *The Theory of Communicative Action: vol.1*, p288.
19. Although Habermas does not exploit the following point, I think a distinction of the type he is making helps explain a kind of inner conflict concerning the expression of speech acts. One might be reluctant, that is, to utter an (intrinsically harmless) illocution for fear of a (threatening, quasi-violent) perlocutionary construal of it. Sensitivity to the perlocutionary force of my avowal of love, for instance, may be such that it overrides the illocutionary value of my uttering it. This may result, alas, in the abandonment of the speech act.
20. *The Theory of Communicative Action, vol.1*, pp292-3.
21. Austin's distinction is introduced in *How to Do Things with Words*, pp98-103. Both Culler and Wood identify this problem with Habermas's argument in the articles cited above, from which I draw here. Wood calls the proposition that the communicative use of language - linguistic interaction oriented towards 'reaching understanding' - is 'originary' the *verständigungs-thesis* ('Habermas's Defence of Rationalism', p154). He cites Habermas's early formulation of it as "fundamental to all speech is the type of action aimed at reaching understanding" (*Communication and the Evolution of Society*, p1), and remarks that the only explicit defence of the thesis follows the appeal to Austin in *The Theory of Communicative Action: vol.1*, pp286-295. While I would agree with this remark, I will later reconstruct an alternative defence of the *verständigungs-thesis* which I will argue is implicit in Habermas's interpretation of Wittgenstein's rule-following considerations.
22. *The Theory of Communicative Action, vol.1*, p293.
23. This point is elaborated by Wood, 'Habermas' Defence of Rationalism', p158.
24. Chris Martin cited this case to me, I trust with illocutionary force bracketed.
25. *The Theory of Communicative Action, vol.1*, p292.

26. Culler, 'Communicative Competence and Normative Force', p136. Habermas has replied to Culler's objection in *Postmetaphysical Thinking*, p86, note 32, and his response does not cover the criticism I am putting here. I will return to the *political* context of illocutionary and perlocutionary forces in my conclusion to the thesis, specifically with respect to propaganda.
27. Wood also makes this point, 'Habermas' Defence of Rationalism', p161.
28. Irreverent and foolhardy perhaps, but I take Culler's article as representative of this position.
29. My source for this kind of objection will be Jean-François Lyotard, *The Postmodern Condition*.
30. The objection is not exclusive to deconstruction; Thompson makes it in his essay 'Universal Pragmatics', in Thompson and Held (eds.) *Habermas; Critical Debates*, pp125-8.
31. Culler, 'Communicative Competence and Normative Force', p137.
32. Habermas, *The Philosophical Discourse of Modernity*, p196.
33. The distinction is made in *Ibid.*, p207.
34. *Ibid.*, p198.
35. *Ibid.*, p199.
36. In other words, by supposing that 'every interpretation is a misinterpretation', the deconstructivist makes it too easy for Habermas to deflect the objection concerning the 'proving of worth' of 'world-disclosive' language.
37. Lyotard, *The Postmodern Condition*, p10.
38. *Ibid.*, p57.
39. *Ibid.*, p16.
40. *Ibid.*, p17.
41. *Ibid.*
42. *Ibid.*, p65.
43. *Ibid.*, p16. Lyotard simply dismisses Habermas's distinction as "superficial". The staged objections I go onto consider are my own reconstruction of the reason behind this dismissal.
44. Dews (ed.), *Autonomy and Solidarity*, Revised edition, p21.
45. Though I should add that this is not a worry which occurs to Lyotard.
46. Habermas, *The Philosophical Discourse of Modernity*, p198.
47. The remarks are most concentrated in Wittgenstein's *Philosophical Investigations*, #143-242, and *Remarks on the Foundations of Mathematics*, Parts I and II.
48. Habermas, *The Theory of Communicative Action*, vol.2, p16.
49. *Ibid.*, p18.

50. Ibid.
51. Ibid.
52. Ibid.
53. Ibid., p17.
54. K.O. Apel, *Towards a Transformation of Philosophy*, p280.
55. Rasmussen's phrase, in *Reading Habermas*, p31.
56. G.P.Baker and P.M.S.Hacker, *Wittgenstein: Rules, Grammar, and Necessity*, p248.
57. Wittgenstein, *Remarks on the Foundations of Mathematics*, p365.
58. "If language is to be a means of communication there must be agreement not only in definitions but also (queer as it may sound) in judgements. This seems to abolish logic but does not do so". Wittgenstein, *Philosophical Investigations*, #242.
59. Wittgenstein, *Remarks on the Foundations of Mathematics*, p323.
60. Tully, 'Wittgenstein and Political Philosophy', *Political Theory*, 17, May 1989, pp172-204.
61. Ibid., p179.
62. Wittgenstein, *Philosophical Investigations*, #289 and #87.
63. Habermas, *The Theory of Communicative Action, vol.1*, p17.
64. This thought suggests that of the two readings of Wittgenstein remarked upon in note 23 of chapter 2, the 'meaning as expression/world-disclosure' view may be better than the 'meaning as use' interpretation.

Chapter Five: Crisis and Critique; Tensions in Habermas's Conception

1. Habermas, *The Philosophical Discourse of Modernity*, p300.
2. Habermas, 'Philosophy as Stand-In and Interpreter', in Baynes et.al. (eds.) *After Philosophy*, p313.
3. Habermas, 'Questions and Counterquestions', in R.Bernstein (ed.) *Habermas and Modernity*, pp210-1.
4. Habermas can therefore be read as a counterpoint to Hobbes. From Habermas's perspective, Hobbes's state of nature represents the self-undermining logic of strategically mediated interaction, but only given a lifeworld which has undergone the differentiations peculiar to modernity. Insisting on the irrevocability of the latter, Habermas does not allow himself the option, taken by Hobbes, of metaphysically grounded laws of nature. For more on the the common ground of contention between Habermas and Hobbes, see Rolf Zimmermann,

- 'equality, political order, and ethics: Hobbes and the systematics of democratic rationality', in Rasmussen (ed). *Universalism vs. Communitarianism*, pp109-128.
5. See Habermas, *Moral Consciousness and Communicative Action*, p201.
 6. See, for example, *The Philosophical Discourse of Modernity*, p29 & p139.
 7. Ibid., p325.
 8. Habermas, 'Philosophy as Stand-In and Interpreter', p313.
 9. Habermas, *The Philosophical Discourse of Modernity*, p315.
 10. Habermas, 'A Reply', in Honneth and Joas (eds.) *Communicative Action*, pp225-6.
 11. Ibid., p226.
 12. Footnote to McCarthy's introduction to Habermas's *The Theory of Communicative Action, vol.1*, p405, note 12. Benhabib also stresses that "formal criteria of rationality should not be conflated with the integrity of forms of life". (*Critique, Norm and Utopia*, p273). But there remains a tension between 'intact intersubjectivity' as the criterion of critique used to mean either a dialogue situation of pure communicative action, or a harmoniously rationalized lifeworld.
 13. Habermas, *The Philosophical Discourse of Modernity*, p324 (my emphasis).
 14. For Habermas's early recognition of these problems, see his 'Some Difficulties in the Attempt to Link Theory and Praxis', in *Theory and Practice*, especially pp22-4. For a lucid summary, see R. Bernstein's introduction to *Habermas and Modernity*, p12f, and Honneth, *The Critique of Power*, pp280-5.
 15. See Habermas, *Moral Consciousness and Communicative Action*, p31; 'Philosophy as Stand-In and Interpreter', p300; *The Philosophical Discourse of Modernity*, p300. Of course, the validity clause does not cover the rationality (efficiency) of instrumental actions.
 16. See Benhabib, *Critique, Norm and Utopia*, pp263-70.
 17. The kind of transcendental argument which is appropriate for showing the necessity of *some* framework for self-interpretation will be considered in the following chapter.
 18. Habermas, *The Philosophical Discourse of Modernity*, p300.
 19. Habermas, 'Philosophy as Stand-In and Interpreter', p314.
 20. Habermas, *Moral Consciousness and Communicative Action*, p196. Habermas retains this view in his more recent writings. For the Kantian, he writes in 'Lawrence Kohlberg and Neo-Aristotelianism' (see following footnote), "What one understands by morality becomes a matter of how one answers the question concerning the possibility of rationally resolving practical questions in general." (p5).

21. Habermas explicitly develops his thoughts on discourse ethics in the essays collected in *Moral Consciousness and Communicative Action*. But this conception of ethics is implicit throughout *The Theory of Communicative Action*, most clearly in his discussion of Mead and Durkheim (vol.2, pp43-113), and in his discussion of the uncoupling of system and lifeworld, pp174-80. Moreover, the basic intuitions underlying discourse ethics can be traced much further back in Habermas's work; to the notion of intact intersubjectivity as the dialogue situation.
- On the other hand, there are also notable discontinuities in the development of Habermas's position. He now prefers to distinguish between the different *uses* of practical reason - moral, pragmatic, and therapeutic - rather than between (moral) practical reason, and other forms of rationality. This new position is outlined in 'On the pragmatic, the ethical, and the moral employments of practical reason', and 'Lawrence Kohlberg and Neo-Aristoelianism', translated by Ciarin Cronin, in *Remarks on Discourse Ethics*, forthcoming. My citations from these articles are from the unpublished manuscripts. While this amendment to Habermas's position deflects some of the criticisms I put in this chapter, I will argue that it fits uneasily with the framework of his theory of communicative competence. Moreover, to the extent that discourse ethics is understood metatheoretically as an exercise in rational reconstruction, my objections hold. Put differently, the changes within Habermas's position in these articles represent a move away from the strong claims for rational reconstructions proposed in his earlier remarks on discourse ethics.
22. Habermas, *Moral Consciousness and Communicative Action*, p197.
23. Nor is the moral point of view determined by the empirical will; "the will determined by moral grounds does not remain external to argumentative reason - the autonomous will becomes completely internal to reason", 'On the pragmatic, the ethical, and the moral employments of practical reason', p13.
24. In contrast to this, there is the argument against the moral/ethical sceptic which appeals to phenomena of crises of identity in bringing the claim to normality to self-clarification. I am proposing that this other (hermeneutic) argument is the more fundamental, and that it is misleading to construe it as of the same rational reconstructive form as 'discourse ethics'.
25. Post-conventional moral consciousness refers to an abstracted, disengaged, cognitive orientation, by which actors put the justification of conventional morality to the test of universalizable principles. Habermas appeals to the developmental moral psychology of Lawrence Kohlberg as empirical support for the claim that it is the final stage of a learning process, typical of the cognitive gains ascribable to modern differentiation and reflexivity.

For a representative account of Habermas's views on this matter, see the title essay of Habermas, *Moral Consciousness and Communicative Action*. For a critique of its conception of advanced moral competence, see Dreyfus and Dreyfus, 'what is morality? a phenomenological account of the development of ethical expertise', in Rasmussen (ed.) *Universalism vs. Communitarianism*, pp237-267.

26. For an excellent discussion of the formal/pragmatic refutation, see Kenneth Baynes, *The Normative Grounds of Social Criticism*, ch3. Also useful is White, *The Recent Work of Jürgen Habermas*, ch.3.
27. Habermas, *Moral Consciousness and Communicative Action*, p104.
28. Habermas, 'Questions and Counterquestions', p210. This is consistent with Habermas' general definition of rationality as criticizability; see *The Theory of Communicative Action*, vol.1; p10.
29. Habermas, *Moral Consciousness and Communicative Action*, p104.
30. Ibid., p108.
31. Ibid., p178.
32. Ibid., p108.
33. Ibid., p109.
34. Ibid.
35. Habermas, 'Questions and Counterquestions', p209.
36. For a discussion of this, see Benhabib, 'In the Shadow of Aristotle and Hegel: Communicative Ethics and Current Controversies in Practical Philosophy', *The Philosophical Forum*, XXI, Fall-Winter 1989-90, especially pp4-13.
37. Habermas, *Moral Consciousness and Communicative Action*, p206; 'Lawrence Kohlberg and Neo-Aristotelianism', p18.
38. Rawls, *A Theory of Justice*, pp423-4. With respect to the problem which follows, Rawls's option of relective equilibrium isn't available to Habermas. He might refer instead to the relevance of norms for social integration, hence dismissing Rawls's fanciful counter-example. While undoubtedly of some force, this move generates its own difficulties, as I go on to indicate.
39. See Benhabib, 'In the Shadow of Aristotle and Hegel: Communicative Ethics and Current Controversies in Practical Philosophy', p12.
40. Habermas, *Moral Consciousness and Communicative Action*, p206.
41. This, for Habermas, is the counterfactual presupposition that participants in argumentation must make.

42. As this remark suggests, Habermas takes aesthetic and therapeutic discourse to be intimately related. The link lies in the 'need interpretative' role of both. Aesthetic discourse is conceived as 'world-disclosive', as against the hypothetical or problem-solving procedures of practical argumentation. But Habermas also links aesthetic discourse to the thematization of the validity claim to authenticity, raising other problems which I go on to discuss.
43. Habermas, *Moral Consciousness and Communicative Action*, p108. I referred to this architectonic in (1.1) above.
44. 'Objective truth' refers here to the validity claim thematized in the discourse of natural science, and fits ill, even within Habermas's cognitive map, with the category 'culture'.
45. See Martin Seel, 'The Two Meanings of Communicative Rationality: Remarks on Habermas's Critique of a Plural Concept of Reason', in Honneth and Joas (eds.), *Communicative Action*, pp36-48.
46. Habermas, 'On the pragmatic, the ethical and the moral employments of practical reason', p4.
47. Ibid.
48. Ibid., pp9-10.
49. Ibid., p10.
50. Ibid. p4.
51. For Taylor's contrasting analysis of the ideal of 'authenticity', see his *The Malaise of Modernity*, pp.25-42.
52. "Cultural values do not appear with a claim to universality, as do norms of action. At most, values are candidates for interpretation under which a circle of those affected can, if occasion arises, describe and normatively regulate a common interest". Habermas, *The Theory of Communicative Action, vol.1*, p20.
53. I follow Benhabib's terminology here; 'In the Shadow of Aristotle and Hegel: Communicative Ethics and Current Controversies in Practical Philosophy', p15.
54. Habermas, *Moral Consciousness and Communicative Action*, p201.
55. Homogeneous, I should add, only in *this* respect. Habermas's concept in other respects is indeed very complex, playing various methodological, epistemological, sociological, and normative roles. Again, this complexity creates its own problems; see for instance those highlighted by Hans Joas in 'The Unhappy Marriage of Hermeneutics and Functionalism', in Honneth and Joas (ed.) *Communicative Action*, especially pp114-8.
56. See Honneth, *The Critique of Power*, pp300-3; and Fraser, 'What's Critical about Critical Theory?', *New German Critique*, 35, for whom the basic issue is not "whether lifeworld

norms will be decisive but, rather, *which* norms will" (p127). In defense of Habermas, this question is secondary to that of the intelligibility of the claim to normativity which Habermas addresses.

57. A similar criticism is put by Alessandro Ferrara, 'Universalisms: procedural, contextual, and prudential', in Rasmussen (ed) *Universalism vs. Communitarianism*, pp11-39. It is a point which recurs throughout Rasmussen's own book *Reading Habermas*.
58. Habermas, *Moral Consciousness and Communicative Action*, p108, & p177-8.
59. Ibid., p206.
60. Though I admit that the extended comment on the relation between justification and application in Habermas's 'A Reply' in Honneth and Joas (eds.), *Communicative Action*, makes my objection seem somewhat beside the point.
61. Habermas, *Moral Consciousness and Communicative Action*, p182.
62. See Benhabib, *Critique, Norm, and Utopia*, pp340-3.
63. For Habermas's reply to Benhabib's objection, see Dews (ed.), *Autonomy and Solidarity*, Revised Edition, pp251-2. His thrust is that in real discourses of application, as opposed to the idealized context of justification, the other must always be concretized. But this response will only seem convincing if one has already accepted the strong distinction between justification and application.
64. Habermas, 'Questions and Counterquestions', p210.
65. Habermas, *Moral Consciousness and Communicative Action*, p104.
66. Ibid., pp177-8.
67. See Habermas, 'Reply to my Critics', pp227-8.
68. Habermas, 'On the pragmatic, the ethical and the moral employments of practical reason', p13.
69. In Dews, (ed.) *Autonomy and Solidarity*, Revised Edition, pp166-7.
70. Ibid., p162.
71. Habermas, *Moral Consciousness and Communicative Action*, p206.
72. "neo-Aristotelian approaches...must demonstrate how an objective moral order can be grounded without recourse to metaphysical premises". Ibid., p214, n.15.
73. For the charge of conservatism, see 'Kohlberg and Neo-Aristotelianism' pp13-14. In a broader political context, Habermas suspects its ideas of deactivating the "explosive force of universalistic principles of morality", thereby minimizing "the burden of moral justification incumbent upon the political system", *The New Conservatism*, pp41-2.

74. This is quite different to the position that the differentiated truth and normative rightness validity claims are always raised *simultaneously*.
75. Habermas, 'Kohlberg and Neo-Aristotelianism', p12.
76. Ibid.
77. Ibid., pp10-11.
78. Ibid., p13.
79. Ibid.

Chapter Six: Taylor's Conception of Crisis

1. For Taylor, some general concept of the person and the good is ineradicable for the understanding and explanation of human affairs, even though particular conceptions of them are contingent to cultures. I am here drawing on the useful distinction between 'concept' and 'conception' made by Putnam in *Reason, Truth, and History*, p117.
2. "We are selves only in that certain issues matter for us": Taylor, *Sources of the Self*, p34. See also Taylor, 'The Concept of a Person', in *Human Agency and Language*, p98. That this is the basic idea upon which Taylor's philosophical position builds is suggested by his 'Reply to de Sousa' in *The Canadian Journal of Philosophy*, 18, 1988, p450.
3. This point is thought to hold for meanings generally. See for instance the analogy drawn in this respect between the meaning of 'courage' and 'red' in Taylor, 'What is Human Agency?', in *Human Agency and Language*, p19.
4. But not vice versa; See Taylor, *The Malaise of Modernity*, pp35-41.
5. Here, as much elsewhere, Taylor draws upon the insights of Heidegger. For the Heideggerian source of this point, see *Being and Time*, Pt 1, Div.1, VI; 'Care as the Being of Dasein'.
6. These axioms of moral intuition are discussed in Taylor, *Sources of the Self*, pp3-25.
7. This raises the question of how the person whose identity (what matters to him) consists in the fulfilment of his de facto desires is to be accounted for as a strong evaluator. It is discussed in (6.3) and (6.4) below, but Nietzsche has given us a snappy answer to the same question arising from the inverse phenomenon: "He who despises himself still nonetheless respects himself as one who despises". (*Beyond Good and Evil*, maxim 78, p74)
8. Taylor, *Sources of the Self*, p63.

9. For Taylor's thoughts on the distinction between first and second order desires, and how his own distinction between strong and weak evaluations relates to it, see Taylor, 'What is Human Agency', *Human Agency and Language*, pp15-26.
10. Taylor, *Sources of the Self*, p93.
11. Ibid., pp95-6.
12. They are mentioned, together with a useful summary of the key terms introduced in *Sources of the Self* in 'Comments and Replies', in *Inquiry*, 34, June 1991, pp237-54, especially p243.
13. This is an extremely brief summary of arguments set out in 'Self-interpreting Animals', *Human Agency and Language*, pp45-76; and 'Interpretation and the Sciences of Man', *Philosophy and the Human Sciences*, pp15-58.
14. Taylor, *Sources of the Self*, p35.
15. Ibid., p36.
16. Ibid., p38. For a much earlier expression of this idea, see Taylor's *Hegel and Modern Society*, p160.
17. Taylor, 'Legitimation Crisis?', in *Philosophy and the Human Sciences*, p287.
18. Compare, for instance, Taylor's remarks in 'What is Human Agency?', where he writes; "The point of introducing the distinction between strong and weak evaluations is to contrast the different kind of self that each involves". (*Human Agency and Language*, p23). This is confusing, since he also claims that it is virtue of strong evaluations that there is such thing as a 'self' at all.
19. Taylor, *Sources of the Self*, p322.
20. Taylor, 'Legitimation Crisis?', *Philosophy and the Human Sciences*, p277.
21. Ibid., pp248ff.
22. Ibid., especially p278f.
23. 'Inwardness and the Culture of Modernity', in McCarthy et al. (eds.) *Zwischenbetrachtungen Im Proceß der Aufklärung*, p622-3.
24. Taylor, *Sources of the Self*, p18.
25. Ibid., p19. I will return to this point when considering Rorty's objections to Taylor's conception of the self in (6.4), below.
26. Taylor, *Sources of the Self*, p27.
27. M. Löw-Beer, 'Living a Life and the Problem of Existential Impossibility', *Inquiry*, 34, June 1991, pp217-36.

28. Ibid, pp224-5. Löw-Beer, however, does not draw attention to the fact that this should be the position of the weak evaluator described by Taylor in 'What is Human Agency', *Human Agency and Language*, pp26-35.
29. Taylor, *Sources of the Self*, p27.
30. Ibid., p28.
31. Ibid., p30.
32. Ibid., p32.
33. Ibid., p30.
34. Ibid., p31.
35. Ibid., p102.
36. As I noted in (1.0), Taylor's most developed discussion of 'neo-Nietzscheanism' is his 'Foucault on Freedom and Truth', in *Philosophy and the Human Sciences*, pp152-85.
37. The 'utilitarian' conception is attacked in 'What is Human Agency', *Human Agency and Language*, pp17f. The 'emotivist' self is MacIntyre's term (see *After Virtue*, p30), and is put to a critique more thorough and subtle than MacIntyre's in *Sources of the Self* ch.19. Taylor criticizes the general view of Rorty's *Philosophy and the Mirror of Nature* in 'Rorty in the Epistemological Tradition', in Malachowski (ed.) *Reading Rorty*, pp257-75, but doesn't deal there with the objections to his own views which can be reconstructed from Rorty's later writings. Löw-Beer mentions Rorty's conception of the self in contrast to Taylor's in 'Living a Life and the Problem of Existential Impossibility', but he does not take up the problems which Rorty's view pose for Taylor's.
38. My sources for Rorty's views are 'Freud and Moral Reflection', in Smith and Kerrigan, (eds.), *Pragmatism's Freud*, pp1-27; and *Contingency, Irony, and Solidarity*, especially ch.2.
39. Rorty, 'Freud and Moral Reflection', p5.
40. Ibid, p12.
41. This is a point of common ground with Taylor, who agrees that the self is always more than the articulations given to it.
42. Rorty, 'Freud and Moral Reflection', p6.
43. Ibid., p18.
44. Ibid., p9; *Contingency, Irony, and Solidarity*, p31.
45. Ibid., p12.
46. Ibid.
47. The phrase is Taylor's, 'Rorty in the Epistemological Tradition', p269.

48. Guignon and Hiley cite evidence of this sort in 'Biting the Bullet: Rorty on Private and Public Morality', in *Reading Rorty*, especially p356, with references on p363. Similar evidence is provided by Stephen Frosh in *Identity Crisis; Modernity, Psychoanalysis and the Self*, eg. p45. Anthony Giddens incorporates it within a sociological perspective in *Modernity and Self-Identity*, especially ch.2.
49. See footnote 10 to 'Freud and Moral Reflection', p22.
50. Wittgenstein, *Tractatus Logico-Philosophicus*, #4.0031.

Chapter Seven: Taylor, Language, and Critique

1. Taylor, *Sources of the Self*, p50.
2. *Ibid.*, p72.
3. Taylor distinguishes between 'ad hominem' and 'apodictic' models of practical reasoning in 'Explanation and Practical Reasoning', *Wider Working Papers WP72*, p2. My citations of this article come with the acknowledgement that it is a working paper circulated in a provisional form to encourage discussion and comment before publication.
4. Taylor, *Sources of the Self*, p72.
5. As Taylor puts it, Nietzsche's "genealogy goes to the heart of the logic of practical reasoning. A hypergood can only be defended through a reading of its genesis". *Sources of the Self*, p73. See also p505.
6. Taylor gives the homely example of the de facto incompatibility of two activities such as football and chess, in contrast to the in principle incompatibility or incommensurability of soccer and rugby football; "the rules which partly define these games prescribe actions in contradiction to each other. Picking up the ball and running with it is against the rules of soccer". *Philosophy and the Human Sciences*, p144. A similar relation holds, he suggests, between pre- and post-Galilean science. Describing nature in anthropocentric terms is a 'foul' in the latter.
7. Taylor, *ibid.*, p149.
8. 'Explanation and Practical Reasoning', pp7-12.
9. *Ibid.*, p12.
10. *Ibid.*, pp11-12.
11. Taylor instances the cruel punishment of a parricide in the mid-18th century as described by Foucault in *Discipline and Punish*, *ibid.*, p13.

12. Ibid., p16.
13. But articulation only becomes an imperative given the modern moral predicament; see *Sources of the Self*, p92f.
14. 'Understanding and Explanation in the *Geisteswissenschaften*, in Holtzman and Leich (eds.), *Wittgenstein: to Follow a Rule*, p200.
15. Ibid., p192.
16. The point that the meaning of virtue terms - or what Taylor calls 'desirability characterizations' - such as courage, integrity, and the like, do not withstand division into 'fact-stating' and 'value-projecting' components, is worth bringing attention to partisans of a 'value-free', 'scientific' reading of Marx. Their efforts to purify Marx's concept of exploitation from evaluative residue, by translating it into a neutral term like 'extraction', 'pumping out' or whatever, fail to pick out something in the world which is essential to what Marx identifies; namely, *exploitative* practices.
17. Ibid., p191. Taylor has an unfortunate penchant for critically engaging with the vulgar cousins of philosophers, especially the distant ones of Marx (eg. *Sources of the Self*, p203).
18. 'Understanding and Explanation in the *Geisteswissenschaften*', p204.
19. Ibid., p205.
20. Taylor, *Sources of the Self*, p69.
21. The orthodox view thus reinforces the 'naturalist' prejudice that first person accounts, articulated in the language of everyday life, are not explanatory.
22. Taylor, *Sources of the Self*, p58.
23. Ibid., p59.
24. The epistemological consideration focusses more on certainty, as against non-conclusive validations of strong evaluation interpretation, and 'essentially contestable' value-terms.
25. Taylor, *Sources of the Self*, p69.
26. Ibid., p57.
27. Ibid., p81.
28. Ibid., p59.
29. Ibid., p67.
30. 'Understanding and Explanation in the *Geisteswissenschaften*', pp205-6.
31. Wittgenstein, *Philosophical Investigations*, #201.
32. Tully, 'Wittgenstein and Political Philosophy', *Political Theory*, May 1989, p 195.
33. See Habermas, 'On Systematically Distorted Communication', p217.
34. See 'A Digression on Historical Explanation', in *Sources of the Self*, pp199-210.

35. Ibid., p207.
36. 'Wholly sympathetic' but only *qua* expressivist. There are important differences, not least her rejection of the absolute conception. But I think my attribute is merited, since that conception is alien to expressivism.
37. Lovibond, *Realism and Imagination in Ethics*, p169.
38. Ibid., p204.
39. Ibid., p173.
40. Ibid., p175.
41. Ibid., p178.
42. Ibid., p194.
43. Ibid., p218.
44. Ibid., p219.

Chapter Eight: The Critique of Reason: Instrumental, Functional, or Procedural?

1. Both Habermas and Taylor express their debt to, and attempt to build upon, the description of the paradoxes of the subject-object relation in the final chapter of Foucault's *The Order of Things*.
2. Taylor, *Sources of the Self*, pp99-100.
3. This is a criticism which has repeatedly been put against Habermas. Taylor makes it in 'Language and Society', p30.
4. Taylor does this to great effect in the chapter on Locke in *Sources of the Self*, 159-74, which I refer to below.
5. A clear, concise, and convincing case for this claim, relating utilitarianism specifically to the dominance of the market in modernity, is made by Ross Poole in *Morality and Modernity*, pp8-17.
6. The *locus classicus* of this view is of course Adorno and Horkheimer's *Dialectic of Enlightenment*. Similar points about utilitarianism (that it leads to the opposite of happiness whether accepted or not) are put less dramatically by Bernard Williams in *Morality*, pp96-112.
7. Habermas, *The Theory of Communicative Action; vol. 1*, pp398-9.
8. Habermas, *The Philosophical Discourse of Modernity*, p29.
9. Ibid., pp315-6.

10. Taylor, *Sources of the Self*, p160
11. Ibid., pp156, 163, 168.
12. It should by now be clear that, for Taylor, the two motives are not ultimately separable. Like Gellner, he thinks that epistemological issues have their root in an ethic of (disengaged) cognition.
13. Taylor, *Sources of the Self*, p171.
14. Ibid., p176.
15. Ibid., p173.
16. Habermas, *The Philosophical Discourse of Modernity*, p297.
17. Taylor, *Sources of the Self*, p.81.
18. Ibid., pp82-3.
19. Ibid., p85.
20. Here Taylor draws on Williams's description of the "institution of morality", which I discussed in (3.1).
21. 'Comments and Replies', *Inquiry* 34, 1991 p252.
22. Though this only comes into focus once Habermas's theory is amended to the satisfaction of critics like Fraser and Honneth, whose objections to Habermas's colonization thesis were noted in chapter five.
23. See Habermas, *The Philosophical Discourse of Modernity*, p139.
24. Taylor, *Sources of the Self*, p510.
25. Ibid., p419.
26. Ibid., pp374, 381, 421.
27. Ibid., p420.
28. Peter Winch puts this expressivist thought as well as anyone; "Reality is not what gives language sense. What is real and what is unreal shows itself in the sense that language has". 'Understanding a Primitive Society', in his *Ethics and Action*, p12.
29. *Sources of the Self*, p428.
30. Ibid., p481.
31. Ibid., pp481-2.
32. Ibid., p481.
33. Ibid., pp482, 425.
34. Ibid., p478.
35. Ibid., p481.
36. Ibid., p465.

37. Ibid., p462.
38. Ibid.
39. Ibid., p480.
40. Ibid.
41. Ibid., p464.
42. Ibid., p480.
43. Ibid., p464.
44. Ibid., p445.
45. Ibid., p448.
46. Ibid., p453.
47. Ibid., p429.
48. Ibid., p472.
49. Ibid., p491.
50. Ibid., p.462.
51. I think this throws some light on Jay Bernstein's insufficiently justified objection against Taylor that by employing "the discourses of modernism directly", he leaves himself open to the charge that "epiphanies are just imagined, mere possibilities without cognitive purchase on us". (Bernstein, *The Fate of Art*, note 6, pp275-6.) In avowed opposition to Taylor, Bernstein believes that "it is not the contents of these [modernist/ epiphanic] works which is most significant but their forms". It is strange that this view should be opposed to Taylor's, since the whole point about the concept of epiphany is that form and content cannot be separated. A more important issue, however, is that the real difficulty for Taylor lies in his reliance on the narrative, substantive model of practical reasoning to deflect the objection that epiphanies are a matter of mere subjective imagination. This is why he deploys discourses of modernism for the sake of content. Only in this way can Taylor address the problem - which to his credit he does not simply wish away - that epiphanies have a cognitive purchase.
52. Ibid., p505.
53. Ibid.
54. Ibid., p428.
55. Ibid., p481.
56. Ibid.
57. And therefore would be open to the same charges which, in (6.4), were put to Rorty from the viewpoint which Taylor defends.

58. A point noted by Michael Rosen, 'Must we Return to Moral Realism?', *Inquiry* 34, June 1991, pp183-94.
59. Taylor, *Sources of the Self*, p515.
60. Ibid.
61. As Bernstein remarks; "psycho-sexual and moral developmental schemes cannot be empirical in this way because they are determined by their terminus, their picture of maturity, autonomy, health, virtue, or the good life for men, which are not themselves subject to empirical questioning" J. Bernstein, 'Narrative and Self-Narration in Psychoanalysis', p64.
62. Taylor, *Sources of the Self*, p513..

Conclusion: The Politics of Democratic Problem-Solving and World-Disclosure

1. Taylor, *Sources of the Self*, p513.
2. Habermas, *Moral Consciousness and Communicative Action*, p211.
3. Ibid.
4. As a philosophical doctrine, discourse ethics seeks to refute skepticism concerning the validity of justice claims. But this is done without considering the role which might be played by *prima facie* moral intuitions of the kind at issue here; they simply don't feature in the rational reconstruction of the moral point of view attempted in discourse ethics.
5. The 'green' intuition that the natural environment and wilderness are bearers of moral claims which human beings ought to recognize requires comment. This intuition might be supported by the simple thought experiment of imagining two worlds, one where there is no life, and another where there is life like there is on earth but without its human form. If we allow that the latter is in some sense a better world than the former, (and not just because there is a greater chance of humans coming into existence), then the challenge might be put to radically rethink the inheritance of anthropocentric grounds of value. While I do not think that the validity of moral intuitions can in general be demonstrated or refuted by appeal to thought experiments, it might be worth considering if the appeal of this particular thought experiment would soon backfire if we were to expand it a little.

In the year 2100 nuclear war destroys all detectable life on Earth. In the year 2050, however, an astronaut had been sent into space in search of life on another planet. After 50 years of Earth-time, she discovers a planet (which she calls 'Tera') which appears to be

life-supporting and very similar to Earth as it is thought to have been just prior to the last Pleistocene Ice Age (very green and without humans) except that it is bigger. Further, one of its moons (which our astronaut calls 'Husk') appears to be as barren - though 'naturally' - as the post-holocaust lifeless Earth. Suddenly, a large meteorite is seen approaching the planet at a very high and deadly speed.

What moral intuitions or judgements would we have or make about these respective heavenly bodies if we were to be that astronaut? Would we want to say that Tera was better, or contained more goodness, than Husk? Would we want to say that Tera was about to become worse than the Earth just prior to the Ice-age, but better than the then contemporary Earth, since the destruction of life on Earth was the result of a human action? Would we want to say anything about the meteorites? Would we want to say anything of moral significance about what caused the Ice-Age?

The sheer obscenity of a radioactively decimated Earth, or even of an Earth depleted of many of its life forms, provokes the 'Deep Ecological' ethical reorientation toward locating value in the life forms themselves, either by pragmatically endowing them with rights which demand respect, or - and this is less embarrassingly humanistic - by infusing them with a mystical quality (Being) which calls for reverence. But our thought experiment indicates that the critique of anthropocentrism may not be quite what it seems. We began by accepting the intuition that Earth minus humans would in some sense be better than an Earth minus life. It is an intuition which seems to be reinforced if we think of the Earth as being destroyed of life by human actions. But if we expand our thinking to consequentially equivalent states of worlds (equivalent in terms of the existence of life and the diversity of species), that intuition no longer seems so applicable. We must be able to differentiate between the impending doom facing the imaginary planet Tera and our imagined fate of the Earth, and not only in terms of the extent of life or species loss; a difference intelligible only if we distinguish between the actions of human beings, and the happenings of nature. What makes a meteorite different to a missile, and the extinction of the brontosaurus different to that of the dodo, is of course the human *responsibility* for these facts. If we focus attention on responsibility, then we can preserve our intuition about the immorality of environmental destruction, and about the need for a proper sense of our inter-dependence with nature, while at the same time recognizing the *political* need for human autonomy as the condition of applying this sensibility. This must, in part, be an autonomy *from* brute nature - as well as from the social 'realm of necessity' - insofar as we are able to bring about happenings for which we can be said to be *at fault* ourselves. It does not follow from this that we must

follow Kant in placing the metaphysical locus of this autonomy and responsibility completely outside nature, since human beings inhabit a world of linguistically disclosed significance which cannot be naturalized to the world inhabited by brutes without remainder. Further, Habermas's approach highlights that the responsibility at stake here is that of a collective 'us'; an 'us' only in the medium of which we forge our identities and the identities of objects with moral claims. One problem with taking solitary aesthetic experience as an exemplary source of moral attunement and self-awareness is that there are so many identity forming conditions presupposed in such experiences, conditions which can only be established on the basis of the *shared* activities of human beings. On the other hand, the fact that human beings do share a linguistically constituted intersubjectivity does not explain the peculiar pull of moral claims once disclosed to a particular subject. What is needed is a conception of responsibility which synthesizes Habermas's and Taylor's insights, though I am not yet in a position to provide it.

6. Taylor, *The Malaise of Modernity*, p91.
7. Habermas, *The Theory of Communicative Action*, vol.2, p.394
8. Heidegger introduces this idea of nature being disclosed as a standing resource (or as the translator puts it, 'standing-reserve'), in 'The Question Concerning Technology', in *The Question Concerning Technology and Other Essays*, translated and with introduction by William Lovitt, p17.
9. I mean 'world picture' here in the sense given to it by Heidegger in 'The Age of the World Picture', *Ibid.*, pp115-154. That essay is not itself, of course, an exercise in deconstruction.
10. A similar view is shared by John Dryzek, who writes that "the human life-support capacity of natural systems is the generalizable interest *par excellence*". *Rational Ecology: Environment and Political Economy*, p36, cited by Robyn Eckersley, 'Habermas and Green Political Thought', *Theory and Society*, 19, 1990, p757. However, I take issue with Dryzek's view (seemingly shared by Eckersley; see *Ibid.* 765-6) that this support-system is best approached as a "partner in communication" in note 17, below.
11. Convincing support for this explanatory hypothesis is given in Timothy W. Luke and Stephen K. White, 'Critical Theory, the Informational Revolution, and an Ecological Path to Modernity', in J.Forrester (ed.), *Critical Theory and Public Life*, pp22-49. The need for a modification to Habermas's theory by way of proposing an institutionalization of communicative norms at the level of production is persuasively urged by Tony Smith, *The Role of Ethics in Social Theory*, pp169-71.
12. Taylor, *Sources of the Self*, p.97.

13. Taylor, *The Malaise of Modernity*, pp81-2.
14. But, for a 'subjectivist' interpretation of Taylor which does leave such sensibilities offended, see Stephen R.L. Clark, 'Taylor's Waking Dream', *Inquiry*, 34, June 1991, pp195-215.
15. For a defence of science as a project oriented to the cooperative use of problem-solving rationality, see Nicholas Maxwell, *From Knowledge to Wisdom*. Although Maxwell is perhaps over ambitious in proposing that "As a result of implementing this Enlightenment programme [qua rational cooperative problem-solving], we may well expect to achieve in life a degree of progress towards what is of value in life to us that is comparable to the remarkable progress that has been achieved in science (in improving knowledge)" (pp156-7), the point to be stressed is that such a project is not to be confused with the accumulation of objectifying knowledge by subjects. Only the latter conception of science undercuts projects of world-disclosure.
16. This suggestion is made by John S. Dryzek in 'Green Reason: Communicative Ethics and the Biosphere', *Environmental Ethics*, 12, Fall 1990, pp195-210.
17. Dryzek writes, "intersubjective discourse presupposes some ecological - and not just linguistic - standards. Although it is easy to forget, our communications with one another can proceed only in and through the media made available by the natural world...if Lovelock is right, the atmosphere in which we live, talk, hear, write, read, smell, and touch is composed and regulated by the planet's biota acting in concert. This biota makes possible and maintains a physical environment fit for itself - and for us, and our communications...Because any such [communicative] act is made possible by this ecological system, it can be called to account in accordance with ecological standards. If indeed nature is a silent participant in every conversation, then perhaps it deserves a measure of the respect that we accord to human participants. If critical communication theorists argue that only entities capable of entering into communication can be assigned value, then there is a sense in which Gaia passes their test." (Ibid., p205).

In this passage, Dryzek shifts from the interesting and eminently plausible idea that there are 'ecological standards' which are presupposed in intersubjective discourse, to the much less plausible suggestion that these presuppositions should be considered as unheard participants in discourse. The latter view is wholly adventitious to the first, and requires the addition of the speculative Gaia hypothesis. The former point can be upheld without relying on such speculation, which in any case betrays a misunderstanding of the nature of communicative ethics. Critical communication theorists need not claim that only beings capable of speech and action are assigned value, but that only such beings, in their capacity

for world-disclosure and rational problem-solving, can meaningfully assign it. Since the procedural test does not function as a mechanism for selecting candidates for participation, it is misleading to propose that Gaia can either pass or fail it.

There is another interesting twist to Dryzek's 'extension' of communicative rationality to include 'ecological standards'. Of the latter, he cites 'diversity, homeostasis, flexibility', which in general are the presuppositions for the maintenance of a 'self-regulating global system' (pp204-5). As such, they are precisely the standards of system maintenance which Habermas is so concerned to keep distinct from the lifeworld. A major difficulty for radical ecology is how to reconcile ascribing value to nature *in toto* with the requirement of acknowledging the moral demands of particulars; a problem Habermas tries to offset by refusing to construe the lifeworld as a meta-subject.

18. Habermas, *The Theory of Communicative Action*, vol.2, pp394-5. This point fits peculiarly well with the suspicion directed by 'rationalist' radical ecologists, such as Dryzek, against those 'Deep Ecologists' who ascribe the environmental crisis to a profane spirituality, or to the suppression of intuitions 'naturally' attuned to the environment. Concerning the latter view - to which, we have seen, Taylor subscribes - Dryzek comments that "interventions in complex systems often have counterintuitive results, as actions ramify throughout these systems. As a result, intuitions, good intentions, and sympathetic sensibilities are insufficient guides to action" (op.cit.p200). In other words, world-disclosure must be constrained by rational problem-solving, even in the case of ecologically attuned action.
19. Since writing this concluding section, I have discovered that in his book *Political Theory and Postmodernism*, Stephen K. White organizes the political debate between defenders of 'modernity' and 'postmodernity' around this distinction between the action-coordinative and world-disclosive aspects of language. Corresponding to this distinction, White proposes, are conceptions of the 'responsibility to act' and the 'responsibility to otherness'; the former is prioritized by defenders of modernity, the latter by postmodernists. I am at one with White in thinking that the political differences between Habermasians and postmodernists can be clarified by appeal to this distinction, and I am in sympathy with his preference for Heidegger's conception of language over post-structuralist conceptions. If anything, White is insufficiently critical of the latter. For a conception of language which is incompatible with the *intelligibility* of communication, as postmodernism's Derrida arguably is, cannot simply be *added on* to a conception which does account for the possibility of communication, just for the sake of its 'world-disclosive' pay-off. Nor is it clear why "concerted deployment of new fictions against whatever fictions are socially in force"

requires "a deep affirmation of the world-disclosing capacity of language" (ibid., p27). What is disclosed by language is the world, not the human power to create fictions. Furthermore, it is hardly a sign of "responsibility to otherness" that I respond to the *truth*-claim of an exploited other with "throws of Nietzschean laughter", as White's view seems to imply (px). Nor does White consider the "cognitive machinery" of ethical/political self-understanding associated with rational action-coordination and world-disclosure in any detail, a lacuna which I believe could be filled by linking it to subjective and objective criticism. But these objections are a little unfair, since they apply to issues outside the remit of his book. This concluding section can be seen as taking up White's offer (ibid. pxii) of applying the distinction between rational action-coordination and world-disclosure to a political topic other than the issue of justice (justice being the topic of his own application of the distinction).

BIBLIOGRAPHY

- Adorno, Theodor and Max Horkheimer. *Dialectic of Enlightenment*, translated by John Cumming, London: Verso, 1986.
- Apel, Karl-Otto. *Towards a Transformation of Philosophy*, translated by G. Adey and D. Frisby, London: Routledge & Kegan Paul, 1980.
- Austin, J.L. *How to Do Things with Words*, 2nd edn., J.O. Urmson and M. Sbisà (eds.), Oxford: Clarendon Press, 1975.
- Baker, G.P. and P.M.S. Hacker, *Wittgenstein: Rules, Grammar and Necessity*, Oxford: Blackwell, 1985.
- Baynes, Kenneth. *The Normative Grounds of Social Criticism*, New York: SUNY Press, 1992.
- Baynes, Kenneth, James Bohman and Thomas McCarthy. (eds.) *After Philosophy: End or Transformation?*, Cambridge MA: MIT Press, 1987.
- Benhabib, Seyla. *Critique, Norm, and Utopia: A Study of the Foundations of Critical Theory*, New York: Columbia University Press, 1986.
- 'Autonomy, Modernity, and Community', in T. McCarthy, C. Offe, and A. Wellmer (eds.), *Zwischenbetrachtungen Im Prozeß der Aufklärung*, (q.v.) pp373-391.
- 'In the Shadow of Aristotle and Hegel: Communicative Ethics and Current Controversies in Practical Philosophy', *The Philosophical Forum*, 21, Fall-Winter 1989-90, pp1-31.
- Bernstein, J.M. 'The Politics of Fulfilment and Transfiguration', *Radical Philosophy*, 47, Autumn 1987, pp21-29.
- 'The Causality of Fate: Modernity and Modernism in Habermas', *Praxis International*, 8, January 1989, pp407-25.
- 'Art against Enlightenment: Adorno's Critique of Habermas', in Andrew Benjamin (ed.), *The Problems of Modernity* pp49-66, London: Routledge, 1989.
- 'Self-Knowledge as Praxis: Narrative and Narration in Psychoanalysis', in Christopher Nash (ed.), *Narrative in Culture*, pp51-77, London: Routledge, 1991.
- The Fate of Art*, Cambridge: Polity, 1992.

- Bernstein, Richard. *Beyond Objectivism and Relativism*, Oxford: Blackwell, 1983.
- (ed.) *Habermas and Modernity*, Cambridge, MA: MIT Press, 1985.
- The New Constellation*, Oxford; Blackwell, 1991.
- Bleicher, Joseph. (ed.) *Contemporary Hermeneutics*, London: Routledge & Kegan Paul, 1982.
- Brand, Arie, *The Force of Reason; An Introduction to Habermas' Theory of Communicative Action*, Sydney: Allen and Unwin, 1990.
- Cascardi, Tony. *The Subject of Modernity*, Cambridge: Cambridge University Press, 1992.
- Cavell, Stanley. *The Claim of Reason*, Oxford: Oxford University Press, 1979.
- Clark, Stephen R.L. 'Taylor's Waking Dream', *Inquiry*, 34, June 1991, pp195-215.
- Culler, Jonathan. 'Communicative Competence and Normative Force', *New German Critique*, 35, Spring-Summer 1985, pp133-44.
- Dallmayr, Fred. 'Life-World and Communicative Action', in *Polis and Praxis: Exercises in Contemporary Political Theory*, pp224-53, Cambridge, MA: MIT Press, 1984.
- 'The Discourse of Modernity: Hegel, Nietzsche, Heidegger (and Habermas)', *Praxis International* 8, January 1989, pp377-406.
- Davidson, Donald. 'Paradoxes of Irrationality', in R. Wollheim and J. Hopkins (eds.), *Philosophical Essays on Freud*, pp289-305, Cambridge: Cambridge University Press, 1982.
- Dews, Peter. *Logics of Disintegration; Post-structuralist Thought and the Claims of Critical Theory*, London: Verso, 1987.
- (ed.) *Autonomy and Solidarity; Interviews with Jürgen Habermas*, Revised Edition, London: Verso, 1992.
- Dreyfus, Hubert L. and Stuart E. Dreyfus 'What is Morality? A Phenomenological Account of the Development of Moral Expertise' in D.Rasmussen (ed.), *Universalism vs. Communitarianism* (q.v.), pp237-266.
- Dryzek, John S. 'Green Reason: Communicative Ethics for the Biosphere', *Environmental Ethics*, 12, 1990, pp195-210

- Eckersley, Robyn. 'Habermas and Green Political Thought', *Theory and Society*, 19, 1990, pp739-776.
- Ferrara, Alessandro. 'Universalisms: Procedural, Contextualist, and Prudential', in D. Rasmussen (ed.), *Universalism vs. Communitarianism* (q.v.), pp11-38.
- Foucault, Michel. 'What is Enlightenment?', translated by Catherine Porter, in Paul Rabinow (ed.), *The Foucault Reader*, pp32-50, Harmondsworth: Penguin, 1986.
- Fraser, Nancy. 'What's Critical about Critical Theory? The Case of Habermas and Gender', *New German Critique*, 35, Spring-Summer 1985, pp97-131.
- Frosh, Stephen. *Identity Crisis; Modernity Psychoanalysis and the Self*, London: Macmillan, 1991.
- Gadamer, Hans-Georg. *Truth and Method*, translated by Garrett Barden and John Cumming, London: Sheed and Ward, 1975.
- Philosophical Hermeneutics*, Berkeley: University of California Press, 1976.
- Gellner, Ernest. *Thought and Change*, London: Weidenfield and Nicolson, 1964.
- Legitimation of Belief*, Cambridge: Cambridge University Press, 1974.
- Spectacles and Predicaments*, Cambridge: Cambridge University Press, 1979.
- Relativism and the Social Sciences*, Cambridge: Cambridge University Press, 1985.
- Culture, Identity, and Politics*, Cambridge: Cambridge University Press, 1987.
- Geuss, Raymond. *The Idea of Critical Theory*, Cambridge: Cambridge University Press, 1981.
- Giddens, Anthony. *Modernity and Self-Identity*, Cambridge: Polity, 1991.
- Glover, Jonathan. *I: The Philosophy and Psychology of Personal Identity*, London: Penguin, 1988.
- Guignon, Charles B. and David R. Hiley, 'Biting the Bullet: Rorty on Private and Public Morality', in A. Malachowski (ed.), *Reading Rorty* (q.v.), pp339-64.
- Habermas, Jürgen. *Toward a Rational Society*, translated by J. Shapiro, London: Heinemann, 1971.
- Knowledge and Human Interests*, translated by J. Shapiro, London: Heinemann, 1972.

- Theory and Practice*, translated by J. Viertel, London: Heinemann, 1974.
- Legitimation Crisis*, translated by T. McCarthy, London: Heinemann, 1976.
- Communication and the Evolution of Society*, translated by T. McCarthy London: Heinemann, 1979.
- 'The Hermeneutic Claim to Universality', in J. Bleicher, (ed.) *Contemporary Hermeneutics* (q.v.), pp181-211.
- 'On Systematically Distorted Communication', *Inquiry*, 13, 1970, pp205-18.
- 'Towards a Theory of Communicative Competence', *Inquiry*, 13, 1970, pp360-75.
- 'Modernity versus Postmodernity', *New German Critique*, 22, 1981, pp3-14 (Reprinted as 'Modernity - An Incomplete Project', in Foster, H. (ed.), *Postmodern Culture*, pp3-15, London: Pluto, 1985).
- 'Reply to my Critics', in John B. Thompson, and David Held (eds.), *Habermas: Critical Debates* (q.v.), pp219-83.
- 'Interpretive Social Science vs. Hermeneuticism', in N. Haan et al. (eds.), *Social Science as Moral Inquiry*, pp251-67, New York: Columbia University Press, 1983.
- The Theory of Communicative Action; vol.1, Reason and the Rationalization of Society*, translated by T. McCarthy, Boston: Beacon Press, 1984.
- 'Questions and Counterquestions', *Praxis International*, 4, 1984 (reprinted in Bernstein, Richard (ed.) *Habermas and Modernity* (q.v.), pp192-216).
- Autonomy and Solidarity; Interviews*, Revised Edition, Peter Dews (ed.), (q.v.).
- The Philosophical Discourse of Modernity: Twelve Lectures*, translated by Frederick Lawrence, Cambridge: Polity, 1987.
- The Theory of Communicative Action; vol 2: Lifeworld and System: A critique of Functionalist Reason*, translated by T. McCarthy, Boston: Beacon Press, 1987.
- Moral Consciousness and Communicative Action*, translated by L. Christian and N.S. Weber, Cambridge MA: MIT Press, 1989.
- The New Conservatism*, edited and translated by S.W. Nicholson, introduction by S. Wolin, Cambridge: Polity, 1989.

- 'A Reply', in Axel Honneth and Hans Joas (eds.) *Communicative Action* (q.v.), pp214-64.
- Postmetaphysical Thinking*, translated by William M. Hohengarten, Cambridge: Polity, 1992.
- 'Lawrence Kohlberg and Neo-Aristotelianism', translated by Ciaran Cronin, in *Remarks on Discourse Ethics*, Cambridge: Polity, forthcoming.
- 'On the pragmatic, the ethical, and the moral employments of practical reasoning', translated by Ciaran Cronin, in *Remarks on Discourse Ethics*, Cambridge: Polity, forthcoming.
- Hegel, G.W.F. *Early Theological Writings*, translated by T.M. Knox, New York: Harper and Brothers, 1961.
- Heidegger, M. *Being and Time*, translated by John Macquarrie and Edmund Robinson, Oxford: Blackwell, 1967.
- The Question Concerning Technology and Other Essays*, translated and with introduction by William Lovitt, New York: Harper and Row, 1977.
- Holtzman, S. and C. Leich. *Wittgenstein: To Follow a Rule*, London: Routledge & Kegan Paul, 1981.
- Holub, Robert C. *Jürgen Habermas: Critic in the Public Sphere*, London: Routledge, 1991.
- Honneth, Axel. *The Critique of Power; Reflexive Stages in a Critical Social Theory*, translated by Kenneth Baynes, Cambridge MA: MIT Press, 1991.
- Honneth, Axel and Hans Joas. (eds.) *Communicative Action*, translated by Jeremy Gaines and Doris L. Jones, Cambridge: Polity, 1991.
- Husserl, E. *The Crisis of European Sciences and Transcendental Phenomenology*, translated by David Carr, Evanston: Northwestern University Press, 1970.
- Jay, Martin. *Marxism and Totality; Adventures of a Concept from Lukacs to Habermas*, Berkeley: University of California Press, 1984.
- 'The Debate over Performative Contradiction: Habermas vs. the Post-Structuralists', in T. McCarthy et.al. (eds.) *Zwischenbetrachtungen im Prozeß der Aufklärung* (q.v.)
- Keat, Russell. *The Politics of Social Theory: Habermas, Freud, and Positivism*, Chicago: University of Chicago Press, 1981.

- Kelly, Michael. 'MacIntyre, Habermas, and Philosophical Ethics', *The Philosophical Forum*, 21, Fall-Winter 1989-90, pp70-93.
- 'The Gadamer-Habermas Debate Revisited: the Question of Ethics', in D. Rasmussen (ed.) *Universalism vs. Communitarianism* (q.v.), pp139-162.
- Kolb, David. *The Critique of Pure Modernity*, Chicago: University of Chicago Press, 1986.
- Kortian, Garbis. *Metacritique*, Cambridge: Cambridge University Press, 1981.
- Lovibond, Sabina. *Realism and Imagination in Ethics*, Oxford: Blackwell, 1983.
- Löw-Beer, Martin. 'Living a Life and the Problem of Existential Impossibility', *Inquiry*, 34, June 1991, pp217-36.
- Luke, Timothy W. and Stephen K. White. 'Critical Theory, the Information Revolution, and an Ecological Path to Modernity', in J. Forrester (ed.), *Critical Theory and Public Life*, pp22-53, Cambridge MA: MIT Press, 1985.
- Lyotard, Jean Francois. *The Postmodern Condition*, translated by Geoff Bennington, Manchester: Manchester University Press, 1984.
- MacIntyre, Alasdair. 'Epistemological Crises, Dramatic Narrative, and the Philosophy of Science', *The Monist*, 60, 1977, pp453-72.
- After Virtue*, London: Duckworth, 1981.
- Three Rival Versions of Moral Enquiry*, London: Duckworth, 1990.
- Malachowski, A. (ed.) *Reading Rorty*, Oxford: Blackwell, 1990.
- Matustik, Martin J. 'Jürgen Habermas at 60', *Philosophy and Social Criticism*, 16, January 1990, pp.61-79.
- Maxwell, Nicholas. *From Knowledge to Wisdom*, Oxford: Blackwell, 1984.
- McCarthy, Thomas. *The Critical Theory of Jürgen Habermas*, Cambridge MA: MIT Press, 1985.
- Ideals and Illusions; On Reconstruction and Deconstruction in Contemporary Critical Theory*, Cambridge MA: MIT Press, 1991.
- McCarthy, T., C. Offe and A. Wellmer. (eds.) *Zwischenbetrachtungen im Prozeß der Aufklärung*, Frankfurt am Main: Suhrkamp, 1989.

- McCumber, John. 'Philosophy as the Heteronomous Centre of Modern Discourse: Jürgen Habermas' in H. Silverman (ed.) *Philosophy and Non-philosophy since Merleau-Ponty*, pp211-31, London: Routledge, 1988.
- McDowell, John. 'Virtue and Reason', *The Monist*, 62, 1979, pp331-350.
- 'The John Locke Lectures', 1991, Unpublished manuscript.
- Nietzsche, F. *The Genealogy of Morals*, translated by Francis Golffing, Doubleday: New York, 1956.
- 'Truth and Falsity in an Ultramoral Sense', in Geoffrey Clive (ed.) *The Philosophy of Nietzsche*, pp503-15, translated by Oscar Levy, New York: Mentor Books, 1965.
- The Will to Power*, Walter Kaufmann (ed.), translated by Walter Kaufmann and R.J. Hollingdale, New York: Random House, 1968.
- Beyond Good and Evil*, translated by R.J. Hollingdale, Harmondsworth, Penguin, 1973.
- Paz, Octavio. *On Poets and Others*, translated by Michael Schmidt, Manchester: Carcanet, 1987.
- Pippin, Robert B. *Modernism as a Philosophical Problem*, Oxford: Blackwell, 1991.
- Poole, Ross. *Morality and Modernity*, London: Routledge, 1991.
- Putnam, Hilary. *Reason, Truth, and History*, Cambridge: Cambridge University Press, 1982.
- Rasmussen, David. *Reading Habermas*, Oxford: Blackwell, 1990.
- (ed.) *Universalism vs. Communitarianism*, Cambridge MA: MIT Press 1990.
- Rawls, John. *A Theory of Justice*, Oxford: Oxford University Press, 1972.
- Redding, Paul. 'Habermas's Theory of Argumentation', *The Journal of Value Inquiry*, 23, 1989, pp15-32.
- Redner, Harry. *The Ends of Philosophy*, London: Croom Helm, 1986.
- Rehg, William. 'Discourse Ethics and the Communitarian Critique of Neo-Kantianism', in *The Philosophical Forum* Vol. 22, Winter 1990, pp120-138.
- Rorty, Richard. *Philosophy and the Mirror of Nature*, Princeton: Princeton University Press, 1979.

'Postmodern Bourgeois Liberalism', *Journal of Philosophy*, 80, 1983, pp583-589.

'Habermas and Lyotard on Postmodernity', in R. Bernstein (ed.) *Habermas and Modernity* (q.v.), pp161-76.

'Freud and Moral Reflection', in J.H. Smith and W. Kerrigan (eds.) *Pragmatism's Freud*, pp1-27, Baltimore and London: John Hopkins University Press, 1986.

Contingency, Irony, and Solidarity, Cambridge: Cambridge University Press, 1989.

Rosen, Michael. 'Must We Return to Moral Realism?', *Inquiry*, 34, June 1991, pp183-94.

Sayre-McCord, Geoffrey. (ed.) *Essays on Moral Realism*, Cornell University Press, 1988.

Smith, Tony. *The Role of Ethics in Social Theory; Essays from a Habermasian Perspective*, New York: SUNY Press, 1991.

Taylor, Charles. *Hegel*, Cambridge: Cambridge University Press, 1975.

Hegel and Modern Society, Cambridge: Cambridge University Press, 1979.

'Understanding in Human Science', *The Review of Metaphysics*, 34, September 1980, pp25-38, followed by discussion with Richard Rorty and Hubert Dreyfus, pp47-55.

'Understanding and Explanation in the *Geisteswissenschaften*', in S. Holtzman and C. Leich, pp191-210, *Wittgenstein: To Follow a Rule*, London: Routledge & Kegan Paul, 1981.

Human Agency and Language: Philosophical Papers Vol.1, Cambridge: Cambridge University Press, 1985.

Philosophy and the Human Sciences: Philosophical Papers Vol.2, Cambridge: Cambridge University Press 1985.

'Overcoming Epistemology', in Baynes, Kenneth et al.(eds.), *After Philosophy: End or Transformation?* (q.v.), pp464-88.

'The Moral Topography of the Self', in S.B. Messer, L.A. Sass, and R.L. Woolfolk (eds.) *Hermeneutics and Psychological Theory*, pp298-320, London: Rutgers University Press, 1988.

'Reply to de Sousa and Davis', *Canadian Journal of Philosophy*, 18, September 1988, pp449-58.

'Inwardness and the Culture of Modernity', McCarthy, T. et al (eds.) *Zwischenbetrachtungen Im Prozeß der Aufklärung* (q.v.), pp.601-23.

Sources of the Self: The Making of the Modern Identity, Cambridge: Cambridge University Press, 1989.

'Explanation and Practical Reason', *Wider Working Papers WP72*, World Institute for Developmental Economics Research, Helsinki, 1989.

'Rorty in the Epistemological Tradition', in A. Malachowski (ed.), *Reading Rorty* (q.v.), pp257-78.

'Language and Society', in Honneth and Joas (eds.) *Communicative Action* (q.v.), pp23-31.

'Comments and Replies', *Inquiry*, 34, June 1991, pp237-54.

The Malaise of Modernity, Ontario: Anansi Press, 1991.

Thompson, John B. *Critical Hermeneutics*, Cambridge: Cambridge University Press, 1981.

Thompson, John B., and David Held. (eds.) *Habermas: Critical Debates*, Cambridge MA: MIT Press, 1982.

Tully, J. 'Wittgenstein and Political Philosophy: Understanding Practices of Critical Reflection', in *Political Theory*, 17, May 1989, pp177-204.

Webern, Anton. *The Path to the New Music*, translated by R. Willi, Bryn Mawr Presser, 1963.

White, S.K. *The Recent Work of Jürgen Habermas*, Cambridge: Cambridge University Press, 1988.

Political Theory and Postmodernism, Cambridge: Cambridge University Press, 1991.

Williams, B. *Morality*, Cambridge: Cambridge University Press, 1976.

Moral Luck, Cambridge: Cambridge University Press, 1981.

Ethics and the Limits of Philosophy, London: Fontana, 1985.

Wittgenstein, L. *Tractatus Logico-Philosophicus*, translated by D.F.Pears and B.F.McGuinness, London: Routledge and Kegan Paul, 1961.

The Blue and Brown Books, 2nd edn. Oxford: Blackwell, 1967.

Philosophical Investigations, translated by G.E.M. Anscombe, 3rd edn. Oxford: Blackwell, 1967.

On Certainty, (ed.) G.E.M. Anscombe and G.H. von Wright, translated by Denis Paul and G.E.M. Anscombe, Oxford: Blackwell, 1969.

Remarks on the Foundations of Mathematics, (ed.) G.H. von Wright, R. Rhees and G.E.M. Anscombe, 3rd edn, Oxford: Blackwell, 1978.

Wood, A.W. 'Habermas' Defense of Rationalism', *New German Critique*, 35, Spring-Summer 1985, pp145-64.

Zimmermann, Rolf. 'Equality, Political Order and Ethics: Hobbes and the Systematics of Democratic Rationality', in D. Rasmussen (ed.) *Universalism vs Communitarianism* (q.v.), pp109-28.