Translating the Sacred: Religion and Postsecularism in the Recent Work of Jürgen Habermas

Dafydd Huw Rees

PhD Social and Political Thought

University of Sussex

March 2015

Contents

Institution: University of Sussex

Name: Dafydd Huw Rees

Degree: PhD Social and Political Thought

Title: *Translating the Sacred: Religion and Postsecularism in the Recent Work of Jürgen Habermas*

Summary

This thesis examines the "postsecular turn" taken by the philosophy of Jürgen Habermas since 2001, with a particular focus on his political theory. It argues that the postsecular turn was motivated primarily by the limitations of Habermas' philosophical paradigm of postmetaphysical thinking. It then analyses his model of postsecular deliberative democracy, and argues that the model should be rejected due to its reliance on an unworkable procedure of sacred-to-secular translation. The thesis is divided into three parts. Part 1 is an analysis of the place of religion in Habermas' writings on social theory from the 1970s and 1980s. It outlines his original account of religious language, and of the "linguistification of the sacred" which accompanies the transition from traditional to modern societies. Part 2 focuses on Habermas' paradigm of postmetaphysical thinking, and shows that the paradigm creates the conditions for postmetaphysical thinkers to appropriate religious concepts. It also argues that Habermas' inability to address the "anthropic problem" in postmetaphysical terms led to his turn to postsecularism. Part 3 examines the model of postsecular deliberative democracy which Habermas has argued for since 2001. Drawing on the accounts of religious language from Part 1 and of appropriation from Part 2, it concludes that the procedure of sacred-to-secular translation on which the model relies is unworkable.

Statement

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Dafydd Huw Rees

10/3/15

Introduction

This is Jürgen Habermas, writing in 1992:

> Under the conditions of postmetaphysical thinking – for which no plausible alternatives exist, despite the fundamentalist reactions against the losses incurred by modernization – the state has lost its sacred substance. This secularization of the spiritual bases of governmental authority, by now long under way, suffers from a deficit in implementation that is overdue. (BFN 443-4)

And this is Jürgen Habermas, writing in 2004:

> The neutrality of state power vis-à-vis different worldviews, which guarantees equal individual liberties for all citizens, is incompatible with the political generalization of a secularized worldview. Secular citizens, in their role as citizens, may neither deny that religious worldviews are in principle capable of truth nor question the right of their devout fellow-citizens to couch their contributions to public discussions in religious language. A liberal political culture can even expect its secular citizens to take part in the efforts to translate relevant contributions from religious language into a publicly intelligible language. (BNR 113)

It is clear that Habermas' view of religion has changed dramatically. His *postsecular turn* is usually dated to his delivery of a paper entitled "Faith and Knowledge," at the Paulskirche in Frankfurt in October 2001. Such a striking change of direction on the part of a major philosopher calls for analysis and explanation. This is what I will offer in this text.

A number of theorists have become associated, in the last few decades, with postsecularism. This text will focus exclusively on Habermas. Rather than juxtaposing his views with those of Charles Taylor[1], José Casanova[2] or others, I will carry out a thorough investigation of Habermasian postsecularism, with the aim of answering several questions. Why did Habermas, late in his career, take such an interest in religion? Why did he develop a postsecular political theory apparently at odds with his earlier work? Does his current position, in fact, reject

---

[1] Charles Taylor, *A Secular Age* (Cambridge, Mass. and London: Belknapp, 2007)

[2] José Casanova, *Public Religions in the Modern World* (Chicago: University of Chicago Press, 1994)

his earlier work or develop it in a new direction? What does Habermas' postsecular turn tell us about philosophy in general, and about political theory in particular? I aim to analyse the postsecular turn and evaluate its results.

"Postsecularism" can mean a number of things, depending on what kind of "secularism" it claims to be "post." There are two broad categories. *Sociological postsecularism* consists of a rejection of the secularisation thesis associated with Max Weber.[3] In contrast to what was once the standard view among sociologists, religious belief does not correlate with low levels of economic development; it is not reasonable, therefore, to expect it to vanish from society over time. This is an empirical matter which I will not deal with in this text. *Political postsecularism* consists of a rejection, or at least a modification, of the principle of political secularism. Religious reasons, argument, and statements should not, in fact, be excluded from politics. This is a matter of political philosophy, and it will be the focus of this text. What interests me about Habermas' postsecularism, then, is the way that he modifies his *discourse theory of law and democracy* in order to construct a model of postsecular politics. The goal of my investigation is to evaluate this model. Before arriving at that point, however, I must locate the postsecular turn in the context of Habermas' philosophical project.

In Part 1 of this text, I look into the original secular standpoint which Habermas has now left behind. Investigating the accounts of *social evolution* and the *linguistification of the sacred* found in his work on social theory from the 1970s and 1980s allows me to clarify the contrast between Habermas' current position and what came before. It illustrates just how dramatic the postsecular turn was. Reviewing Habermas' writings on social theory also allows me to reconstruct his account of religious language, as a "syndrome" of validity claims, attitudes and world-relations. I make use of this account in Part 3.

---

[3] See Max Weber, *The Protestant Ethic and the Spirit of Capitalism* (London: Routledge, 2001). Weber in fact understands modernity as a process of rationalisation, and speaks of "disenchantment" (*Entzauberung*) rather than secularization.

In Part 2, I investigate the complex relationship between religion and the paradigm of *postmetaphysical thinking*, which Habermas outlines in various texts from the 1980s and 1990s. As its name suggests, this is an approach to philosophy which refrains from making metaphysical claims. Ironically, the metaphysical abstinence of Habermas' paradigm of philosophy creates the conditions for non-believers to a*ppropriate* concepts and ethical values from religious discourses. I make use of this account of the appropriation of religion by postmetaphysical thinkers in Part 3. Postmetaphysical thinking is also the key to explaining the postsecular turn. I will argue that it was the inability of Habermas' paradigm to deal with what I call the a*nthropic problem* which pushed him in a postsecular direction. The latter half of Part 2, then, contains my attempt to answer the question of why Habermas made his postsecular turn.

In Part 3, I move on to the political consequences of the postsecular turn, and evaluate the model of *postsecular deliberative democracy* which Habermas has argued for since 2001. After reviewing his criticisms of secular models of deliberative democracy, I investigate the fulcrum of his model, the procedure of *sacred-to-secular translation*. I evaluate this procedure using the accounts of religious language and postmetaphysical appropriation outlined in the first two parts of the text. I argue that the translation procedure is unworkable – it would be a species of *pyrrhic translation*, unable to perform the function which Habermas ascribes to it. On this basis, I conclude that the model of postsecular deliberative democracy does not succeed.

The body of this text focuses closely on Habermas' work. In the Conclusion, I use my findings about Habermas' postmetaphysical paradigm and his postsecular political model to address broader questions. If I am correct about the connection between the postsecular turn and postmetaphysical thinking, what does this tell us about attempts by philosophers to refrain from metaphysics? If I am correct about the strengths and weaknesses of Habermas' political model, what are the implications for political theories which attempt to reconcile the diversity of beliefs present in modern societies with the requirements of deliberative democracy?

Habermas' postsecular turn is a particular moment in the thought of a particular philosopher. But the things we can learn by studying it – about philosophy, about democracy, about belief – are profound.

List of Abbreviations

Texts by Jürgen Habermas

AWM   *An Awareness of What is Missing*

BNR   *Between Naturalism and Religion*

BFN   *Between Facts and Norms*

CES   *Communication and the Evolution of Society*

EFP   *Europe the Faltering Project*

FHN   *The Future of Human Nature*

HE   "History and Evolution"

JA   *Justification and Application*

LC   *Legitimation Crisis*

LG   "The Language Game of Responsible Agency and the Problem of Free Will: How can epistemic dualism be reconciled with ontological monism?"

MCCA   *Moral Consciousness and Communicative Action*

P   "The Political - The Rational Sense of a Questionable Inheritance of Political Theology"

PT   *Postmetaphysical Thinking*

T   "Transcendence from Within, Transcendence in this World"

TCAI   *The Theory of Communicative Action I*

TCAII *The Theory of Communicative Action II*

TJ   *Truth and Justification*

TP   *Theory and Practice*

Texts by John Rawls

IPPR   "The Idea of Public Reason Revisited"

PL   *Political Liberalism*

Part 1

Linguistifying the Sacred

And the Lord said, Indeed the people are one and they all have one language, and this is what they begin to do; now nothing that they propose to do will be withheld from them.

Genesis 11:6

I am afraid that we have not got rid of God because we still have faith in grammar...

Friedrich Nietzsche, *Twilight of the Idols*

1.1 Introduction – Religion in Social Evolution

In this first part of the text, I will analyse the role of religion in Habermas' writings on social theory from the 1970s and 1980s – texts such as *Communication and the Evolution of Society, Legitimation Crisis* and his magnum opus, *The Theory of Communicative Action*. My goal is to isolate Habermas' original account of religious thought and language, so as to use it in my evaluation of his postsecular theories in Part 3. The other function of Part 1 is to clarify Habermas' original view of religion's place in society, before the postsecular turn of 2001. This will enable us to see what kind of "secularism" Habermas is today attempting to leave behind.

I interpret Habermas' theory of social evolution using the metaphor of a *centrifuge*. A centrifuge is a piece of apparatus which is used in industry and scientific research for separating fluid suspensions into their component elements. It consists of a drum, fitted with slots for holding vials of the suspension, attached to a rotor. When switched on, the drum spins rapidly, generating centrifugal force which pushes the heavier elements of the suspension towards the base of the vial. The fluid which is removed from the centrifuge at the end of the process – the supernatant – has been separated into its component elements, which appear as layers of sediment and liquid in the vial. What a centrifuge does, then, is to differentiate what was once

entangled. In my metaphor, the components of the social formation which Habermas describes – *organizational principles*, *institutional cores*, and *forms of social integration* – are the drum of the centrifuge. The process of social evolution is the equivalent of the spinning of the drum. Language and thought are the equivalent of the suspension which is loaded into the centrifuge. As we will see, Habermas believes that language, like a suspension, begins as an undifferentiated mass, in which distinct elements – *validity claims*, *attitudes*, and *world-concepts* – are tangled together. Social evolution forces these elements to differentiate; at the same time, it is the differentiation of these elements – the growth in the communicative capacity of human beings – which makes social evolution possible. Each stage in the process of the differentiation of language correlates to a particular type of society and a particular *worldview*. The *religious-metaphysical* worldview which is my focus is an intermediate stage in this process, halfway between the undifferentiated language of *myth* and the wholly differentiated language of *modernity*. By peering into the centrifuge halfway through its cycle, we can isolate Habermas' theory of religious thought and language; by examining the supernatant which emerges at the end, we can clarify his original account of religion's place in modern society.

I will begin with a section explaining what Habermas aims to achieve with his theory of social evolution: a communicative reconstruction of Marx's historical materialism (1.2). I will then set out the components into which he divides each social formation (1.3), and explain how he thinks social evolution takes place (1.5). The upshot of this investigation is that there are three historical stages in Habermas' theory, namely the *primitive, traditional a*nd *modern* social formations, each of which has a corresponding worldview (1.6). The last three sections of Part 1 (1.8-10) explore these worldviews in detail. Along the way, it will be necessary to clarify Habermas' accounts of the connection between communication and morality (1.4) and his formal pragmatic view of language and communicative competence (1.7).

1.2 Reconstructing Historical Materialism

Habermas calls his theory of social evolution a *reconstruction of historical materialism.* [4] In *Communication and the Evolution of Society*, he briefly describes the process of reconstructing a theory as "taking (the) theory apart and putting it back together again in a new form in order to attain more fully the goal it has set for itself." (CES 95) But Habermas is doing more than tweaking Marx's theory in order to give it greater predictive power or factual accuracy, as others have done. He is applying the method of *rational reconstruction* to the theory, in order to bring to light some of its neglected aspects.

Rational reconstruction is one of Habermas' most important ideas. He developed the method in texts such as in *On the Logic of the Social Sciences*[5] and the "Postscript" to *Knowledge and Human Interests,*[6] and later made it a central tenet of his postmetaphysical approach to philosophy. (PT 7) It is meant to lie halfway between the empiricist or positivist methodology of the natural sciences and the interpretive or hermeneutic methodology of the humanities. Habermas' criticisms of positivism[7] and hermeneutics[8] are well known. Rational reconstruction

---

[4] Jürgen Habermas, *Zur Rekonstruktion des Historischen Materialismus* (Frankfurt am Main: Suhrkamp, 1976). Most of the essays in this volume were translated into English as *Communication and the Evolution of Society*, with the exception of the important article "Geschichte und Evolution." I will be using the English version, "History and Evolution," which appeared in *Telos* 39 (1979), pp 5-44.

[5] Jürgen Habermas, *On the Logic of the Social Sciences* (Cambridge: Polity, 1990).

[6] Jürgen Habermas, *Knowledge and Human Interests* (London: Heinemann, 1981), see "Postscript," pp 351-386. In this Postscript, Habermas refines his analysis in the main text, and distinguishes the "reconstructive" from the "emancipatory" sciences. He had previously conflated them. The emancipatory sciences are those which de-reify the social or subjective worlds, thus creating the possibility for radical change. Psychoanalysis and critical theory itself are examples. Reconstruction overlaps with this to some extent, but Habermas thinks that there are enough differences to justify a separation.

[7] See Jürgen Habermas, "The Analytical Theory of Science and Dialectics" pp 131-162, "A Positivistically Bisected Rationalism" pp 198-225, in Theodor W. Adorno ed., *The Positivist Dispute in German Sociology* (London: Heinemann, 1976).

[8] See Alan R. How, *The Habermas-Gadamer Debate and the Nature of the Social: Back to Bedrock* (Aldershot: Avebury, 1995).

is supposed to avoid the problems of these two methods, and to be an appropriate methodology for postmetaphysical philosophy and critical social theory. Habermas cites Noam Chomsky's linguistics, Jean Piaget's developmental psychology and Lawrence Kohlberg's moral psychology as examples of the reconstructive method.[9]

To rationally reconstruct something means, firstly, to make *explicit* what was previously *implicitly* known. It turns "know how" into "know that,"[10] pretheoretical into theoretical knowledge. (CES 9) It is a commonplace that human beings always know far more than they are aware of knowing. Carrying out a reconstruction of a mundane practice reveals the often intricate capacities and stores of knowledge which lie behind it. To give a linguistic example, a reconstruction would not focus on the explicitly known content of a speaker's utterance, but on the implicitly known rules which the speaker uses in making the utterance. As Jørgen Pedersen puts it, the aim of rational reconstruction is not "a direct paraphrase or translation of an unclear meaning, but rather underlying rules and structures as conditions for any meaningful linguistic expression."[11] Similarly, Thomas McCarthy says that the goal of the process is "not a paraphrase or translation of an originally unclear meaning but an explicit knowledge of rules and structures, the mastery of which underlies the competence of a subject to generate meaningful expressions."[12]

Secondly, rational reconstruction makes explicit *general* structures and *universal* capacities. The method is not appropriate for just any stores of implicit knowledge, for example the implicit knowledge behind purely local phenomena and activities. Habermas says that "reconstructions relate to pretheoretical knowledge of a general sort, to *universal capabilities*,

---

[9] See Jørgen Pedersen, "Habermas' Method: Rational Reconstruction," *Philosophy of the Social Sciences* 38:4 (2008), p 465.

[10] See Gilbert Ryle, "Knowing How and Knowing That: The Presidential Address," *Proceedings of the Aristotelian Society*, New Series, vol. 46 (1945).

[11] Pedersen op. cit., p 462.

[12] Thomas McCarthy, *Ideals and Illusions* (Cambridge, Mass.: MIT Press, 1991), p 131.

and not only to particular competences of individual groups… or to the abilities of particular individuals." (CES 14) A rational reconstruction of the English language would uncover not only the deep structures of this particular language and the implicit knowledge possessed by English speakers, but of language as such and of language-users as such. A rational reconstruction of the game of chess would uncover not only the rules of chess, but the capacity for rule-following, as such, possessed by all players of all games. Habermas says that "*All* species competences of subjects capable of speech and action are accessible to a rational reconstruction if, namely, we recur to the practical knowledge to which we intuitively lay claim in tried-and-true productive accomplishments." (PT 14) The objects of rational reconstruction are universal in the sense that they are *specie-wide capacities*. They are common to all normally-developed human beings.

Thirdly, reconstruction has what can be called a *diachronic* or *developmental* aspect. More than simply uncovering the general structures behind the capacities of speaking and acting subjects, reconstruction tells us about how those general structures change over time – how they emerge, how they develop and how they are acquired by subjects. A rational reconstruction of a phenomenon will reveal its internal history. To use Pedersen's terminology, the reconstruction "uncovers some fundamental competencies (horizontal reconstruction), but also the way these competencies have developed over time (vertical reconstruction.)"[13] Reconstructing a language, for example, should reveal the developmental stages through which it has passed, and the way in which it is acquired by new speakers.[14]

It is clear why Habermas considers Chomsky's linguistics to be an example of rational reconstruction. (CES 14) Chomsky sees "grammar" as an implicit ability or store of knowledge which is innate in all normal human beings, enabling them to learn and master languages. The

---

[13] Ibid, p 463.

[14] Other points could be added to the definition. Pedersen emphasises the fact that rational reconstruction can be applied only to lifeworld phenomena (p 461), and that the hypotheses it generates must be falsifiable (pp 463-4).

theory purports to explain the capacity for language acquisition possessed, uniquely, by human infants.[15] Habermas considers his own theories to be based on rational reconstructions. Formal pragmatics reconstructs the human capacity for communication, making it explicit in the procedures for redeeming validity claims. (CES 26) The discourse theory of morality reconstructs the procedures of moral argumentation in which we engage, making them explicit and formal in the (D) and (U) principles. (MCCA 65-6) The discourse theory of law and democracy, according to Habermas, aims to "rationally reconstruct the self-understanding of… modern legal orders," (BFN 82) making them explicit and formal in the democracy principle. (BFN 110)

Habermas' theory of social evolution "reconstructs" historical materialism in the sense that it applies to the theory knowledge gained by rationally reconstructing the communicative and normative capacities of human beings.[16] By doing so, it corrects what Habermas sees as a flaw in Marx's theory: its over-reliance on instrumental reason. Consider Marx's most well-known summary of historical materialism, in the "Preface" to *A Contribution to the Critique of Political Economy.* Marx argues that certain *forces of production* (labour power, resources, methods of organization, technology) correspond to certain *relations of production* (laws, types of property, political structures), and that the latter are dependent on the former: "The mode of production of material life conditions the social, political and intellectual life-process generally."[17] This allows Marx to identify several types of society, based on their economic structure, namely (in chronological order) the "Asiatic, ancient, feudal and modern bourgeois modes of production."[18] It also gives him a theory of crisis and social evolution. As Marx puts it,

---

[15] See Neil Smith, *Chomsky: Ideas and Ideals* (Cambridge: Cambridge University Press, 1999), especially chapter 2, "The Linguistic Foundation."

[16] This is my interpretation of what Habermas is doing. Markus Patberg, however, argues that the sense of "reconstruction" in Habermas' reconstruction of historical materialism is entirely different to the full sense of "rational reconstruction." See his "Supranational Constitutional Politics and the Method of Rational Reconstruction," *Philosophy and Social Criticism* 40:6 (2014), p 503.

[17] "*A Contribution to the Critique of Political Economy* – Preface," in Karl Marx, *Later Political Writings*, ed. Terrell Carver (Cambridge: Cambridge University Press, 1996), p 160.

[18] Ibid p 160.

> At a certain level of their development the material productive forces of society come into contradiction with the already existing relations of production, or in what is a merely legal expression for this, with the property relations within which they had previously functioned. From forms of development of the productive forces these relations turn into their fetters. Then an epoch of social revolution commences. With the alteration of the economic foundations the whole colossal superstructure is more or less rapidly transformed.[19]

In this model the development of the forces of production is the motor of social evolution. When these forces outpace the relations of production, the relations must change, and it is at this point that a society can be said to have qualitatively changed – in Marx's terminology, to have advanced to a new mode of production. Now, all these examples of the forces of production – technology, industrial methods, the organization of labour power – exemplify what Habermas calls instrumental reason: unreflexive, means-end rationality. (TCAII 333) This is what Habermas objects to. As Michael Schmid puts it, he criticises Marx's theory on the basis that it "placed too much emphasis on economic structures in the sense of purely instrumental action structures. Habermas regards the restriction of a theory of evolution to structures and rules of instrumental action as unjustified and indeed superfluous…"[20] According to Habermas, "the rationalization of action takes effect not only on productive forces but also, and independently, on normative structures." (CES 117) He claims, in contrast to Marx, that learning processes take place

> in the dimension of moral insight, practical knowledge, communicative action, and the consensual regulation of action conflicts – learning processes that are deposited in more mature forms of social integration, in new *productive relations*, and that in turn first make possible the introduction of new productive forces. (CES 98-9)

It is not the development of the forces of production alone which moves history. By applying the findings of rational reconstructions of the normative and communicative aspects of human behaviour to historical materialism, Habermas intends to re-balance the theory, showing that social evolution is driven by advances in communicative as well as instrumental reason.

---

[19] Ibid p 160.

[20] Michael Schmid, "Habermas's Theory of Social Evolution," in John B. Thompson and David Held ed., *Habermas: Critical Debates* (London: Macmillan Press, 1982), p 167.

This distinction between *instrumental* and *communicative reason* is central to Habermas' thought. Instrumental reason – a concept which originated with Max Weber and was adopted by the first generation of Frankfurt School theorists[21] – is the kind of rationality which is merely concerned with finding the most efficient means to achieve a given end, without reflecting upon the value of that end. It is objectivating, and aims to understand the world in order to control it. It is embodied in technology, market transactions, and the exercise of power. While instrumental reason gives human beings greater control over their environment, it also inflicts this domination on individuals and the social world. Habermas considers theorists such as Weber, Adorno and Horkheimer to have placed too much emphasis on the growth of instrumental reason. Their fixation on this aspect of reason, to the exclusion of all others, accounts for the pessimistic nature of their theories of modernity.

Habermas attempts to redress the balance in his own social theory, by emphasising the importance of communicative reason. This is the kind of rationality employed in situations where interlocutors aim to co-ordinate their actions by achieving an unforced mutual understanding (TCAI 101). It is cooperative and intersubjective, and is able to reflect on the worthiness of ends as well as the suitability of means. Communicative reason is exemplified in c*ommunicative action*. To anticipate some concepts which I will clarify in section 1.4, communicative actions are those which can only succeed if all the parties involved come to an unforced mutual agreement about a validity claim. Communicative reason and action are at home in the subset of society which Habermas (following Husserl) calls the *lifeworld*, while instrumental reason governs the *system*, which in a modern society means the state and the economy. (TCAII 153-5) Of course, Habermas is not oblivious to the harmful tendencies of instrumental reason, as his theory of the colonization of the lifeworld by the system indicates.

---

[21] See Thomas McCarthy, "Reflections on Rationalization in *The Theory of Communicative Action,*" in Richard J. Bernstein ed., *Habermas and Modernity* (Cambridge, MA: MIT Press, 1988), p 176.

(TCAII 355-6) He simply wishes to redress the balance of the major theories of modernization – those of Marx, Weber, Adorno and Horkheimer – by telling the communicative side of the story.[22] If modernity means the growth of communicative as well as instrumental reason, it contains a potential for emancipation as well as oppression.

Habermas' reconstruction of historical materialism, then, is part of this broader attempt to inject communicative reason into his predecessors' theories. With these considerations in mind, let us turn to Habermas' theory of social evolution.

---

[22] "As Habermas's theory of action distinguishes non-instrumental action structures, he can easily extend the scope of an evolutionary theory to include *these* structures of communicative and moral action." Schmid op. cit., p 167.

1.3 The Social Formation

Marx divides the mode of production into two components, the means and relations of production. He identifies three major modes of production in history, the ancient, feudal and capitalist. Habermas divides the social formation into three components: *organizational principles, institutional cores,* and most importantly forms of *social integration*. He identifies three major types of social formation: *primitive* (*vorhochkulturell*), *traditional* (*hochkulturell*), and *modern-capitalist*.[23] Each social formation has a prevalent belief system or worldview which is correlated with the level of communicative and moral development of its form of integration, as I will explain in sections 1.8-10. Suffice it to say, for now, that primitive societies have *mythical* worldviews, traditional societies have *religious-metaphysical* ones and modern societies have *modern* or *posttraditional* worldviews. (CES 103-6) The crucial difference between Habermas' theory and that of Marx is the presence of the form of social integration. This component of the social formation is the locus of the changes in communicative and normative capacity which he wishes to illuminate.

Let us consider these components one by one. Habermas describes the organizational principle as an abstract set of rules which characterizes a social formation. (LC 7) It delimits how, and to what extent, a society can change without becoming a different kind of society. (CES 153) As Habermas puts it, the organizational principle "determines especially within what structures changes of the institutional and interpretation system are possible; in what quantity existing productive power capacities can be socially used or the development of productive power itself can be stimulated; and thus also how far the steering achievement, i.e. the system-complexity of a society, can be increased." (HE 31) Habermas seems to conceive of organizational principles as the fundamental practical values of a society ("how should we do things?"), which colour all

---

[23] Habermas sometimes draws finer distinctions, for example between "liberal" and "organized or advanced" capitalist societies, and between both of these and "post-capitalist class societies" (meaning presumably the planned economies of the 20th century). He also mentions a further stage, "post-modern society," but says little about it. (LC 17)

of its practices and provide a framework for dealing with any new challenge. Habermas offers several examples. Age and sex roles are the organizational principle of primitive societies (LC 18). Political class domination (LC 18-9), otherwise described as the vassal/lord relationship of feudalism (HE 38), is the organizational principle of traditional societies. The wage-labour/capital relationship is the organizational principle of capitalist societies (LC 20-1, HE 38). These are all abstract relationships, rather than stated codes of principles or concrete institutions.[24]

Organizational principles are manifest in institutional cores (or, in some passages, "core institutions"). The institutional core is the most important institution within a society. It seems, judging by Habermas' examples, to have the functions of stabilizing its society, co-ordinating action within it, and ensuring the society's reproduction. A primitive society has a kinship system for its institutional core. (LC 18) Traditional societies have the state (LC 19), and capitalist societies have the market economy (with a complementary and enabling state). (LC 20-1, HE 38) Institutional cores are concrete institutions within their societies, rather than abstract relationships. They are also pervasive – there is no getting away from the market in a capitalist society, or the kinship system in a tribal society. Although Habermas thinks that organizational principles are in some way manifested or embodied in institutional cores, this is not a strict one-to-one relationship. Organizational principles are so abstract that they are compatible with several different institutional cores (CES 153), and the actual embodiment which arises in a given society will be influenced by that society's history.

The form of social integration is the third component of the social formation, and the fulcrum of Habermas' theory of social evolution. Its function is to socialize individual subjects

---

[24] Schmid criticises the explanatory power of the concept on this basis: "organizational principles are categorically defined only as 'abstract rules,' the essential characteristic of which is that they delimit the ranges of variation for structural changes. This definition is clearly insufficient for *differentiating* various types of social formations… We should not assume that a certain social formation can be comprehensively described in terms of an organizational principle alone." Schmid op. cit., p 171.

and coordinate their actions. For individuals to work together and function in society, they must be able to agree to follow rules and recognise common legitimate norms. The form of a society's social integration determines how they do so – how they utilise their capacities for communicative action and normative agreement. Evidently, a form of social integration is a diffuse thing, which cannot be identified in any particular location, in the way that the wage-labour/capital relationship can be localised in the market. To the extent that he can localise it, Habermas associates it with a society's mechanisms for socializing individuals, turning them into subjects able to interact and organize themselves, communicatively or otherwise. Habermas identifies forms of social integration in "structures of interaction" (CES 148) or "normative structures," and associates them with "mechanisms for regulating conflict," "world views" and "identity formations." (CES 120) Forms of social integration involve "knowledge of a moral-practical sort and not technically useful knowledge that can be implemented in rules of instrumental or strategic action" (CES 146) – the element which Habermas thinks is missing from Marx's theory. Habermas gives some examples of social phenomena associated with social integration, including the "family, forms of state organization, universalistic exchange and legal traffic with corresponding collective identities, such as tribe, empire, nation, etc." (HE 32) He asserts that the form of social integration makes possible "the social use of productive forces, the formation of new forms of society, new media etc." (HE 32) Another way of understanding this is via the analogy Habermas makes between social integration and technology in *Legitimation Crisis*. Just as technology allows a society to extend its steering capacity over outer nature – the objective world – processes of socialization allow it to extend its steering capacity over inner nature – the subjective world. (LC 9) Both phenomena involve discourse centred on a validity claim, respectively truth and moral rightness. (LC 9-10) Forms of social integration can

be seen as the technology of inner nature. Innovation with regard to this technology can alter the way a society operates just as much as alternations in the technology of outer nature.[25]

Taken together, these three components enable the social formation to persist and develop. Habermas says that every new mode of production "means a new form of social integration, which crystallizes around a new institutional core." (CES 147) He asserts that "organizational principles of society can be characterised... through the institutional core that determines the dominant form of social integration" (CES 154), and that "the introduction of a new principle of organization means the establishment of a new level of social integration." (CES 160) The crucial point, for Habermas, is that it is changes in the form of social integration – changes to communicative and moral capacities – that *enable the other two components to develop*, and thus enable the social formation to evolve. He argues that "the development of these normative structures is the pacemaker of social evolution, for new principles of social organization mean new forms of social integration; and the latter, in turn, first make it possible to implement available productive forces or to generate new ones, as well as making possible a

---

[25] In later texts, Habermas draws a distinction between *social* and *system integration*: "the former attaches to action orientations, while the latter reaches right through them. In one case, the action system is integrated through consensus, whether normatively guaranteed or communicatively achieved; in the other case it is integrated through the nonnormative steering of individual decisions not subjectively coordinated." (TCAII 150) Social integration works in a normative and communicative way, while system integration works in a strategic way. The coordination of behaviour through communicative action within the lifeworld is the clearest example of social integration; the coordination of behaviour via monetary incentives in the market is the clearest example of system integration. (TCAII 150) The former uses the symbolic medium of language, while the latter uses the non-symbolic steering media of money and power. (TCAI 324) This is certainly an important distinction - Habermas' account of the colonization of the lifeworld turns on the replacement of social by system integration, and the penetration of non-symbolic steering media into the lifeworld. However, when analysing Habermas' theory of social evolution, it make sense to bracket social and system integration together. They fulfil the same function – coordinating the actions of individuals – albeit in different ways. The uncoupling of system from social integration is a product of social evolution. In primitive societies, which have yet to differentiate themselves into system and lifeworld aspects, they are one and the same. (TCAII 153-4, 163-4) Finally, since Habermas holds that communicative action is the basic form of human action on which strategic action is parasitic (TCA1 287-8), it is reasonable to interpret system integration as reliant on social integration. Subjects must be socialised, must possess some communicative and normative competence, before they can be coordinated by system imperatives. At the very least, they must know what it means to obey a command or take part in a transaction. For these reasons, in my exposition of Habermas' theory of social evolution I consider the form of social integration to include both social and system integration.

heightening of social complexity." (CES 120) As we will see in section 1.5, Habermas thinks that it is exogenous events which *trigger* the crises leading to social evolution, but that it is the degree of normative and communicative capacity possessed by a society's form of social integration which determines how it *responds* to these crises. What Habermas is saying, as I interpret it, is that the core institution is the machine; the organizational principle is the blueprint for the machine; and the form of integration is the training and knowledge possessed by the operator, without which the machine could not function at all. Figure 1 summarizes what we have learned about the theory of social evolution so far.[26]

Social Formation Component	Primitive	Traditional	Modern
Organizational Principle	Age and sex roles	Political class domination	Wage-labour / capital relationship
Institutional Core	Kinship system	State	Market
Form of Social Integration	?	?	?
Prevalent Worldview	Mythical	Religious - metaphysical	Modern, posttraditional

Figure 1. Habermas' model of society.

We have seen how the centrifuge is constructed. We have yet to see how it spins. Before moving on to Habermas' account of the process of social evolution, however, a blank area on the diagram must be filled in. What characterises the form of social integration in each social formation? This component, as I have noted, is the locus of advances in communicative and normative capacity. To describe the nature of social integration at each stage, then, I must examine Habermas' view of communication and morality, and the connection between the two.

---

[26] Compare my interpretation to that of David Ingram, *Habermas: Introduction and Analysis* (Ithaca and London: Cornell University Press, 2010), pp 312-6, especially Table 8, p 313. Ingram focuses on organizational principles and says little about the other components.

1.4 Language and Morality

A form of social integration can be at one of three levels of communicative competence: symbolically mediated interaction, propositionally differentiated speech, or an intermediate stage between the two. Habermas says that communicative competence can be seen as a capacity of individual speaking and acting subjects, "or as concepts of the infrastructure of the action system itself. I would like to use them in this latter sense to characterize different forms of social integration." (CES 155-6)

The earliest and most primitive level of communication is *symbolically mediated interaction* (a term borrowed from George Herbert Mead), which Habermas also calls the holistic mode of speech. Habermas argues that in this kind of communication "representation, expression, and behavioural expectation are one." (CES 36-7) At this stage, the propositional and illocutionary components of speech acts are inseparable. Propositional content is the information which a speech act contains, whereas illocutionary force is the effect the speech act has.[27] (CES 36-7) In a speech act such as saying "Fire," the propositional content can be used in a number of illocutionary acts, for example a warning, an order, a request and so on. Habermas asserts that at the level of symbolically mediated interaction, speakers cannot use the same propositional content to perform different illocutionary acts. He also thinks that it is not possible at this stage for a speaker to separate herself from her performative attitude, i.e. she is only able to take a participant's stance on her interaction, not that of an observer. (CES 136-7) This lack of separation is the feature which he associates most strongly with symbolically mediated interaction, noting that at this level "actions, motives (or behavioural expectations), and acting subjects are perceived on a single plane of reality." (CES 155) Most importantly, Habermas thinks that speakers at this stage of communicative competence cannot distinguish between the three

---

[27] In this analysis Habermas draws on and modifies the speech act theory of J. L. Austin and John Searle. See Austin, *How To Do Things With Words* (Oxford: Clarendon, 1962), and Searle, *Speech Acts: An Essay in the Philosophy of Language* (Cambridge: Cambridge University Press, 1969)

world-concepts, the three basic attitudes and the three validity claims. (TCAII 54) This is a crucial point, which I will explain in section 1.7.

Symbolically mediated interaction eventually gives way to *propositionally differentiated speech, w*hich Habermas also calls argumentative speech.[28] At this level, propositional and illocutionary components can be separated. Habermas considers this "uncoupling" to be a fundamental achievement of human language. Speakers become aware that they are communicating on two levels simultaneously – that of content and that of intersubjective relationships and roles. (CES 41-2) It is now possible for speakers to separate themselves from their own performative roles and adopt each other's, as well as adopting an observer's perspective on their interaction. All this has the effect of separating actions from norms, so that the former are seen to be in need of justification by the latter. (CES 155) Habermas says that this level of communicative competence involves "communications that require passing from communicative action to discourse." (CES 82) Speakers can now separate actions, norms and principles. Norms (and roles) are seen to be in need of justification as well, and can be justified with reference to principles. (CES 155) Habermas asserts that at this level of communication, "the validity claims we connect with speech acts can be made thematic. In grounding assertions or justifying actions in discourse, we treat statements or norms (underlying the actions) hypothetically, that is, in such a way that they might or might not be the case, that they might be legitimate or illegitimate." (CES 155) Speakers acquire the ability to distinguish between the three world-relations and basic attitudes, and thus are able to raise and contest validity claims. This mature level of communicative competence therefore enables speakers to move to what Habermas calls discourse.

---

[28] Habermas occasionally describes argumentative speech as a third stage, following propositionally differentiated speech, (CES 155) but by and large he contrasts both of these to symbolically mediated interaction.

Discourse, for Habermas, means the kind of communication in which interlocutors can identify and challenge validity claims. He argues that there are three of these:

> Those claims are claims to truth, claims to rightness, and claims to truthfulness, according to whether the speaker refers to something in the objective world (as the totality of existing states of affairs), to something in the shared social world (as the totality of the legitimately regulated interpersonal relationships of a social group), or to something in his own subjective world (as the totality of experiences to which one has privileged access). (MCCA 58)

If I say "Calcium carbonate is an alkali," I am asserting something about the external world. I am raising the validity claim of truth, a claim that my utterance accurately describes reality. My interlocutor can challenge this validity claim by questioning whether the utterance is in fact valid for the external world – for example, by asking for empirical evidence.[29] If I say "I am happy with my job," I am asserting something about my subjective world. I am raising the validity claim of truthfulness or sincerity, a claim that my utterance is a valid expression of my subjective state. Again, my interlocutor can challenge this claim. This is less straightforward than challenging the claim to truth, since I have privileged access to my subjective feelings and beliefs. But challenges to sincerity are undoubtedly possible. My interlocutor might point out inconsistencies between my statement and previous statements and actions, which could be seen as revealing my subjective state – for example, being late for work, expressing dislike for my co-workers, or looking for another job.

If I say "we should not enslave people," I am asserting something about the normative relations which exist between people in our shared intersubjective world. This is the most complex of the validity claims, and the one to which Habermas has devoted most attention throughout his writings. His discourse theory of morality elaborates on the procedures which must be used to redeem the validity claim of rightness. To say that a rightness claim corresponds

---

[29] The question of how the validity claim of truth is to be redeemed – in other words, what Habermas' theory of truth actually is – is complex. Habermas seem to have moved away from the "consensus theory of truth" suggested by his early text "Wahrheitstheorien," which has anti-realist and relativist implications, towards the position outlined in *Truth and Justification.* At the same time, he decisively rejects the correspondence theory of truth on the basis that it is too metaphysical. (TJ 216-7)

to the intersubjective world does not mean correspondence to the norms of society *as they currently stand* – clearly, 250 years ago most people believed that slavery was acceptable. We would not want to say that slavery was therefore *not* morally wrong at the time. In the passage above, Habermas is careful to say that it is correspondence to the "*legitimately* regulated interpersonal relationships," i.e. those which are regulated by valid norms, which count. He proposes two principles for assessing the validity of norms. The principle of universalization (U) states that:

> *All* affected can accept the consequences and the side effects [the norm's] *general* observance can be anticipated to have for the satisfaction of *everyone's* interests (and these consequences are preferred to those of known alternative possibilities for regulation). (MCCA 65)

The principle of discourse (D) states that:

> Only those norms can claim to be valid that meet (or could meet) with the approval of all affected in their capacity *as participants in a practical discourse*. (MCCA 66)

So according to Habermas, when I say that "we should not enslave people," what I mean is that my statement conforms to (U) and (D) – it would be agreed to under ideal conditions of discourse[30] by all affected parties in the long run. (TCAI 99, MCCA 197) My interlocutor can challenge this claim by questioning whether my statement would in fact pass the tests of (U) and (D).

These, then, are the three validity claims.[31] Habermas says that competent modern language-users can thematize (isolate and challenge) them in the rarefied form of communication he calls discourse. In everyday communicative action the three claims are raised simultaneously and regarded unproblematically. However, should they encounter a problem

---

[30] This would be moral discourse, one of three varieties of practical discourse. The others are pragmatic and ethical discourse. See Habermas' essay "On the Pragmatic, The Ethical, and the Moral Employments of Practical Reason." (JA 1-17) I will deal with the distinction between moral and ethical discourse in § 2.4.

[31] This is my attempt to clarify how validity claims work. It is notoriously difficult to find a straightforward definition anywhere in Habermas' writings. See Joseph Heath, "What is a Validity Claim?" *Philosophy and Social Criticism* 24:23 (1998).

such as a disagreement or a breakdown in communication, competent speakers can shift from communicative action to discourse, where these problems can be solved through a careful testing of validity claims. This option is not available to language-users at the level of symbolically mediated interaction, who do not have the communicative capacity to thematize validity claims. So from Habermas' point of view, there is an internal connection between communicative and normative competence: only language-users who have mastered propositionally differentiated speech are capable of engaging in moral discourse, and thus of generating legitimate moral norms and principles. Morality is not a universal, ahistorical constant, for Habermas. It evolves with language.

What goes for individuals, of course, also goes for components of the social formation. Forms of social integration at the lower levels of communicative competence will not involve interactions which can switch to discourse, while those at higher levels will. (CES 82) The ranking of forms of social integration according to their level of communicative competence is part of Habermas' attempt to reconstruct historical materialism. Given Habermas' view of the internal connection between communication and morality, to say that a society's social integration displays a certain level of communicative competence is simultaneously to say that it is at a certain level of moral development. Habermas insists that measures of moral development apply to forms of social integration as well as the individuals within them: "If one examines social institutions and the action competences of socialized individuals for general characteristics, one encounters the same structures of consciousness." (CES 98-9)[32]

---

[32] In other words, Habermas takes a theory of individual development (ontogenesis) and uses it as a model for social development (phylogenesis). This has attracted considerable criticism. Piet Strydom attacks Habermas for committing the "ontogenetic fallacy" – see Piet Strydom, "The Ontogenetic Fallacy: The Immanent Critique of Habermas' Developmental Logical Theory of Evolution." *Theory, Culture and Society* 9 (1992). Strydom considers any application of an ontogenetic theory to a phylogenetic process to simply be mistaken. Klaus Eder makes a more nuanced criticism, of his own early work on social evolution as well as Habermas'. He argues that if social evolution is to be understood as a learning process (which I will discuss below), then we must base our model of it on a theory of collective, not individual learning. See Klaus Eder, *The Social Construction of Nature* (London: Sage, 1996), and Piet Strydom, "Sociocultural Evolution or the Social Evolution of Practical Reason?:

Habermas classifies these normative structures, or levels of moral development, using concepts borrowed from the developmental psychologist Lawrence Kohlberg. There are three broad stages in Kohlberg's schema. At the *preconventional* stage, norms are obeyed because of sanctions or rewards. Imagine saying to a child "Do X, because I say so." Kohlberg divides this stage into two orientations. Within the "punishments and obedience" orientation, the subject considers good or right actions to be those which lead to the avoidance of punishment, and obeys authority simply because of its ability to impose sanctions. Within the "instrumental relativist" orientation, the subject considers right actions to be those which lead to the satisfaction of its needs. It can guy the system to get what it wants, in a manner reminiscent of what Habermas calls strategic action.[33]

At the *conventional* stage, norms are obeyed because they are prevalent in the group to which the subject belongs. Imagine saying to a child "Do X, because it is what we are expected to do." Again, Kohlberg divides this stage into two orientations. Within the "interpersonal concordance" orientation, the subject considers behaviour which meets with general approval to be good. The behaviour of the majority is reified and considered "natural." Within the "law and order" orientation, dutiful behaviour which maintains the status quo is valued. In a rather Hobbesian manner authority is respected for its own sake, regardless of how it is constituted or what it does.[34]

Eder's Critique of Habermas," *Praxis International* 13:3 (1993). In other words Habermas should have reached for Vygotsky, or someone like him, rather than Piaget and Kohlberg. Habermas seems to have been aware of this problem, since he insists that the relationship between the two theories is a loose one. He asserts that his theory of phylogenesis uses Piaget's account of ontogenesis as a mere "homology," rather than a full analogy. (CES 102-3)

[33] See Lawrence Kohlberg, "From Is To Ought: How to Commit the Naturalistic Fallacy and Get Away with It in the Study of Moral Development," in Theodore Mischel ed., *Cognitive Development and Epistemology* (New York/London: Academic Press, 1971), p 164.

[34] Ibid, p 164.

At the *postconventional* stage, norms are obeyed because they can be justified by legitimate principles or reasoned arguments. Authorities and the standards of groups are themselves held to account. Kohlberg also calls it the autonomous or principled level. Imagine an adolescent replying "Why *should* I do X?" The first orientation within this stage is the "social-contract legalistic" orientation. There is a concern with individual rights and revisable laws which have been agreed to by all. Procedural rules and consensus are important. Kohlberg associates this orientation with the American constitution and legal system. This is followed by the "universal ethical principle" orientation. Autonomy and universality become important at this level – principles which justify actions or norms should be self-chosen, applicable to all, and abstract (like the Golden Rule) rather than concrete (like the Ten Commandments). For Kohlberg, Kantian morality and the Categorical Imperative exemplify this highest level of moral development.[35] Habermas does not want moral development to end with Kant. He is critical of Kant's moral philosophy on the basis that it is monological – individuals test maxims for their universal applicability *in foro interno* (JA 48-9, 60) – and part of the paradigm of the philosophy of consciousness. (PT 124) For these reasons Habermas posits a third postconventional orientation, in which moral rightness emerges from "the communally followed *procedure* of redeeming normative validity claims discursively." (CES 90) He sees his own discourse theory of morality, which operates in this public and dialogical way, as a postmetaphysical advance on Kantian moral theory.

As with communicative competence, these levels of moral development apply to forms of social integration as well as individuals. (CES 98-9) Habermas says explicitly that social integrations can be at one of the three levels of communicative competence (CES 155-6) and of moral development. (CES 99) And, as we have seen, Habermas thinks that there is an internal connection between communicative competence and moral development. He associates

---

[35] Ibid pp 164-5.

symbolically mediated interaction with preconventional morality, the intermediate stages of language with conventional morality, and propositionally differentiated speech with postconventional morality. In his theory of social evolution, then, social formations can be classified in terms of the levels of communicative capacity and moral development embodied in their forms of social integration. A society can be ranked as preconventional and capable of symbolically mediated interaction, and so on. Figure 2 incorporates this information into my model of Habermas' theory.

The form of social integration of the traditional social formation is split between symbolically mediated interaction and propositionally differentiated speech. We will see in section 1.9 why this is the case, why Habermas thinks that this split in communication grounds

Social Formation / Component of Society	Primitive	Traditional		Modern	
Organizational Principle	Age and sex roles	Political class domination		Wage-labour/ capital relationship	
Institutional Core	Kinship system	State		Market	
Form of Social Integration	Symbolically mediated interaction	Symbolically mediated interaction	Propositionally differentiated speech	Propositionally differentiated speech	Level of Communicative Competence
	Pre-conventional	Conventional		Post-conventional	Level of Moral Development
Prevalent Worldview	Mythical	Religious-metaphysical		Modern, post-traditional	

Figure 2. Habermas' model of society – expanded.

31

conventional morality, and what this has to do with his account of religious thought and language. For now, we can move on to the details of social evolution. The picture I have sketched so far of the social formations is static. It is time to spin the centrifuge, and see how, according to Habermas, one social formation develops into the next.

1.5 Crisis and Learning

Habermas believes that social evolution happens through societal *learning* as a response to *crises*. When a social formation "learns," it rearranges its components (organizational principle, core institution, and form of social integration) in such a way that it can overcome the crisis which confronts it. In this way, the primitive social formation evolves into the traditional and the traditional into the modern.

Habermas counts any event which challenges the *steering capacity* of a society as a problem. By "steering capacity," a term taken from systems theory[36], he means the ability of a system to manage complexity. (HE 32) These problems arise continually. The amount of steering capacity a social formation has is closely connected to its organizational principle (LC 7), since it is the organizational principle which determines how much a society can change. (CES 153) Problems which cannot be dealt with using the steering capacity currently available to a social formation count as crises. (CES 160) Unlike mundane problems, they are existential threats to the social formation. Habermas says that "it is external change that overloads the narrowly limited steering capacity of societies organized along kinship lines and undermines the familial and tribal identities." (LC 18) In other words, primitive social formations face exogenous crises, such as demographic change, ecological damage or invasion. (CES 162) Traditional social formations face endogenous crises, centred on conflicts over inequality, leading to class struggles and violence, (LC 20) and modern social formations face legitimation crises. The form of social integration, which is crucial for mobilising the other components (CES 120), determines how much flexibility and steering capacity a social formation has; the level of communicative and normative capacity embodied in this component, then, is an important factor in the outcome of a crisis. As we will see in the following sections, a society's social integration is also

---

[36] See for example Niklas Luhmann, *The Differentiation of Society* (New York: Columbia University Press, 1982). Much of Habermas' social theory from the 1970s and 1980s can be seen as a response to Luhmann's variety of systems theory. The article "History and Evolution" is a direct reply to an article of the same name by Luhmann.

connected to its prevalent worldview, and worldviews play a role in managing and defusing crises. Forms of social integration are therefore all-important in evolutionary crises.

The outcome of a crisis, says Habermas, is the product of two things: the events which cause it, and the degree of adaptability (level of communicative and normative capacity) possessed by the form of social integration. (HE 31) Only the latter factor can be rationally reconstructed. Rationally reconstructing the species-wide capacities embodied in a form of social integration reveals that it is governed by a *developmental logic*. (HE 31) Social integration *has to* pass through the stage of conventional morality before it can reach postconventional morality, and it is possible for an observer to reconstruct this process logically. On the other hand, the events which trigger crises are contingent and unpredictable, and can only be described in terms of *developmental dynamics*. There is no inevitable pattern in this case. Habermas thinks that social evolution is a result of the interplay of the reconstructable developmental logic of forms of social integration and the contingent developmental dynamics of events: it depends on "the developmental logic of the pattern of previous structures and [on] *an impulse given by problem-generating events*." (HE 31)[37]

It is important for Habermas' theory that he only gives an account of the developmental-logical side. Developmental dynamics is an empirical matter, to be investigated by historians using the methods of social science. Developmental logic, on the other hand, may be investigated by postmetaphysical philosophers using the method of rational reconstruction.[38] If

---

[37] On this point see David Owen, "Habermas' Developmental Logical Thesis: Universal or Eurocentric?" *Philosophy Today* 42:1 (1998).

[38] This strict division between developmental logic and dynamics can cause problems. The (reconstructed) developmental logic of forms of social integration is at the heart of Habermas' theory of social evolution. At times, especially in "History and Evolution," he seems to want to remove his theory entirely from the realm of empirical testing: "A narrative always deals with particular events; essentially individual-constants appear in it. A generic competence (*Gattungskompetenz*), however, which lays claim to universality does not even fall under the category of an event that could attract historical attention. Thus, the emergence of such a competence is only accidentally connected with identifiable persons and groups." (HE 18) And again: "sociological considerations are not at all called for here, because they fall short of the level of abstraction on which the structural conditions of possibility of learning processes relevant to evolution must be given." (HE 28) The problem here is that two elements of Habermas'

he was to find a predictable pattern in events themselves, and not just the way that social formations can respond to events, Habermas would be retreating from his postmetaphysical commitments, to a metaphysical "philosophy of history" on Hegelian or Marxist lines. Distinguishing between developmental logic and dynamics, and insisting that only the latter triggers social evolution, allows Habermas to retain the basic materialism of Marx's approach. He denies that he is "quietly dropping the materialist assumptions regarding the motor of social development" and explaining "the dynamics of species history… through an internal history of the spirit." (CES 120-1)

Habermas says that societies "learn" as a response to crises. This is a twofold process. A social formation will first attempt to overcome the crisis using all the "structural possibilities" which its current configuration of components allows. Recall that a society's organizational principle sets a limit to how much it can change, and how much steering capacity it has at its disposal. Habermas would refer to this as staying within the same learning level. When this does not succeed, the social formation must make "one of the improbable evolutionary leaps that lead to *new* learning levels." (HE 32) What this means, in effect, is that the social formation adopts a new set of components, which possess a greater degree of steering capacity and are able to overcome the crisis. The most important component in these evolutionary leaps is the form of social integration. Habermas says in several places that in order to overcome system crises and attain new learning levels, a society must alter its social integration. (CES 122, HE 32) When Habermas says that a society "learns," he means that its form of integration advances to a higher level of communicative competence and moral development, thus allowing it to adopt a new organizational principle and core institution in order to overcome a crisis.

---

postmetaphysical approach to philosophy – his commitment to falsifiable theories, and his refusal to engage in metaphysical philosophy of history – are pulling in different directions. See Pedersen op. cit., p 480, and § 2.3 below.

Habermas believes that in any society, there will be individuals and marginal groups who have achieved higher levels of moral and communicative competence than the social formation as a whole. (CES 121-3) In a primitive society they would be capable of conventional morality and some propositionally differentiated speech, for example. Over time these "learning achievements" filter through into the society's stock of unthematized background knowledge. They circulate in the lifeworld, or are depicted in worldviews. Here they lie latent, waiting to be utilised on a societal scale. They are "structures of consciousness that are already clearly established in interpretive systems but have not yet found institutional embodiment in action systems." (CES 157) The advanced communicative or moral capacity is a "surplus," a "cognitive potential that can be used socially." (CES 160) In the face of crises small groups, which Habermas refers to ambiguously as *social movements*, transpose these latent capacities out of the lifeworld and into institutions. "The rationality structures that find expression in world views, moral representations, and identity formations (…) become practically effective in social movements, and are finally embodied in institutional systems…" (CES 98)[39] More advanced moral and communicative capacities affect the social formation, in the first place, by changing its form of social integration; this allows it to adopt a new organizational principle, and in turn a new core institution, with greater steering capacity. These small groups wield such influence by virtue of having a solution to the crisis which threatens society. Habermas does not think that they consciously aim at implementing higher levels of morality and communication in institutions;

---

[39]Axel Honneth and Hans Joas have criticized Habermas for his overly functional understanding of social movements. They argue that Habermas turns them into mere mechanisms for fixing a social formation's problems. In fact, "social movements do not "learn" in reaction to system-threatening dangers, but in the collective experiencing of, and the co-operative opposition to, injustice arising out of the nature of the social system… (c)onsciousness of social repression and social injustice is not the same as the perception of unsolved problems in the social system." Axel Honneth and Hans Joas, *Social Action and Human Nature* (Cambridge: Cambridge University Press, 1988), p 166. Habermas' use of the ambiguous term "social movements" may be to blame here. He seems to be using it in the same way as the sociologist Alain Tourraine. (CES 125) Honneth and Joas, however, are interpreting it in the everyday sense, as one might speak of feminist or national liberation movements. Even putting this matter of interpretation aside, social movements are an ambiguous element in Habermas' theory. He himself is critical of Shmuel Eisenstadt's ideas about "elites and counter-elites" being the "bearers of innovative potentials" (HE 27-8), but what else are social movements?

they may simply enact these new capacities in their everyday practices. By doing so, the social movements demonstrate how effective the advanced levels of communication and morality are for co-ordinating action, solving conflicts, procuring legitimation, and so on. Some examples of transitions between social formations will clarify what Habermas means.

The first concerns the transition from the primitive to the traditional social formation, which Habermas describes in "Towards a Reconstruction of Historical Materialism." Habermas distinguishes between the "structures of action" which regulate the resolution of conflicts, and those which regulate normal, conflict-free situations. (CES 156-7) The former, in a primitive society, will be preconventional; in other words, the primitive social formation will have a preconventional "legal system." It would be a system of "arbitration and feuding law" (CES 156) and its judgements would be "bound… to the contingent constellations of power of the involved parties." (CES 161) Settlements of conflicts and punishments for crimes vary with a person's status; there is no "rule of law." In contrast, conventional morality predominates in conflict-free situations. Habermas attributes this to the fact that the norms of the kinship system govern these mundane situations. (CES 157) His reasoning, presumably, is that since the kinship system is an extended family, individuals can be persuaded to obey the general rules of behaviour which are approved of by their kin – "do this, because it is what people of our lineage should do" – and not simply bribed or threatened into conforming. This is what Kohlberg calls the "interpersonal concordance orientation," the earliest part of the conventional stage. Habermas also thinks that conventional morality can be found in the mythical worldviews which prevail in this social formation. As he puts it, "in many myths of primitive societies there are already narratively constructed models of conflicts and their resolutions that correspond to the conventional stage of development of moral consciousness; at the same time the institutionalized law satisfies the criteria of the preconventional stage of moral consciousness." (CES 157) People tell stories of gods and spirits resolving conflicts according to conventional norms, even as their actual legal institutions dispense preconventional justice. So in this case, even though the social formation

is at the level of symbolically mediated interaction and preconventional morality, the capacity

for conventional morality (and presumably some level of propositionally differentiated speech)

is latent in mythical worldviews and in the lifeworld.

Habermas suggests that some "evolutionarily promising" (CES 162) primitive societies

would already have established a legal subsystem structured on conventional lines. Such a

conventional subsystem would abstract away from the status of conflicting parties, and judge

all offences against the normative consensus of the kin group – a dogmatic and unquestionable

consensus, it is true, but one which would apply equally to everyone. The promising primitive

society which Habermas imagines might select temporary judges to administer justice at the

conventional level, for those cases which were felt to be ill-suited to the preconventional legal

system. Power would accrue to these officials, as well as a new kind of legitimacy, based on their

function of administering justice in a conventional manner. This is in contrast to the legitimacy

of a tribal leader, which is based on their place in the kinship system, i.e. their descent from a

mythological common ancestor. (CES 161)

This social formation would eventually find itself faced with a crisis "which could not be

managed with an adaptive capacity limited by the kinship principle of organization." (CES 162)

Habermas suggests that environmental problems caused by overpopulation or conflicts over

economic inequality[40] could constitute such a crisis. The preconventional legal system would be

unable to process these problems, and so the conventional subsystem would be increasingly

called upon to solve them.[41] As Habermas puts it, a "few societies under the pressure of

---

[40] This seems to contradict Habermas' earlier assertion that primitive societies only face exogenous crises, and that conflicts over inequality begin in traditional societies. See LC 20.

[41] To quote Ingram: "The decentralized authority that chiefs hold within archaic (tribal) society is limited in dealing with social conflicts and permanent crises calling for more centralized and powerful means. Its reliance on *preconventional* instrumental reciprocity (exchange of favours/disfavours) and fear of punishment, such as ostracism, ignores distinctions between civil and criminal offences, intentions and consequences, and provides few resources for resolving feuds beyond negotiated settlements and ostracism of losing parties… The *state principle of organization* can be seen as an evolutionary solution to these problems. Emerging at the dawn of urbanization and civilization, this principle of organization

evolutionary challenges from such problems made use of the cognitive potential of their

worldviews and institutionalized – at first on a trial basis – an administration of justice at a

conventional level." (CES 162) The temporary officials who administered conventional justice

would begin to take on the status of permanent rulers, in an ever-more hierarchical society, and

the state would replace the kinship system as the society's core institution. This institutional

change would expand the social formation's steering capacity, and allow it to overcome its crisis.

Habermas adds that appropriate conditions of stabilization – such as military victory over

neighbouring tribes (CES 162) – would be necessary for the new social formation to establish

itself. So in this example, the officials who act as temporary judges within the conventional legal

subsystem count as a "social movement." They bring the normative structure of conventional

morality out of its latency in worldviews, and embody it in institutions. The society has attained

a more advanced form of integration (conventional morality and some level of propositionally

differentiated speech), which allows it to form a new core institution (the state) and

organizational principle (political class domination). The primitive has evolved into the

traditional social formation.

The second example concerns the transition from the traditional to the modern social

formation, which Habermas describes in "History and Evolution." He takes as his starting point

Benjamin Nelson's account of the universalistic ways of thinking which first emerged in Europe

in the 12th and 13th centuries.[42] Nelson argues that the new concepts which appeared at this

time – "*universitas, civitas, communitas, persona, libertas, conscientia, liber, machina"* and so

on (HE 36) – enabled the eventual shift away from the probabilistic thinking characteristic of the

early Middle Ages, to a more precise understanding of nature and of human behaviour, in which

---

depends on *conventional* legal norms rather than personal prestige in maintaining order. Such norms
transcend personal familial authority and its dependence on simple threats in resolving disputes and
enforcing compliance." Ingram op. cit., p 315.

[42] See Benjamin Nelson, *On the Roads to Modernity: Conscience, Science, and Civilizations – Selected
Writings* ed. Toby E. Huff (Totowa, New Jersey: Rowman and Littlefield, 1981).

phenomena are subsumed under universal categories. It was this new way of thinking, according to Nelson, which made possible the cognitive advances of the scientific revolution and the Reformation during the 16th and 17th centuries. What characterises these new concepts, says Habermas, was "universally oriented thinking based only on arguments." (HE 36) This way of thinking was discovered, lay dormant in worldviews, and was finally embodied in institutions several centuries later. However, Habermas is dissatisfied with Nelson's Eurocentrism, and suggests that these universalistic concepts in fact emerged during what Karl Jaspers called the Axial Age, between 800 and 200 BCE. (HE 37-8) During this period, according to Jaspers, new ways of understanding the world emerged spontaneously in Greece, Israel, India and China. The great religions and philosophies of this era are the products of this advance, which was guided by figures such as Plato, the Buddha, and Confucius.[43] Habermas sees this as a process of rationalizing mythological worldviews. "These doctrines overcome mythical thinking to the extent that they objectify the world as a whole, distinguish the natural order from the historical one, develop the concept of an abstract law, trace the multiplicity of phenomena to principles, and replace narrative explanations by argumentative ones... individualized concepts of the self and soul also appear." (HE 37-8) This, as far as Habermas is concerned, was the actual origin of the nomological, discursive and universal ways of thinking which were utilized during the transition to the modern age.

---

[43] Karl Jaspers, *The Origin and Goal of History* (London: Routledge and Keegan Paul, 1953). The Axial Age is one of Habermas' favourite references. He uses frequently in his recent writings on postsecularism, which are beset by a phenomenon which could be called "axial inflation" As part of his postsecular paradigm, Habermas asserts that many modern ways of thinking have religious roots. He often does this by insisting that they originate in the Axial Age. (FHN 108, BFN 141-2, 246, AWM 17) In order to make this debatable claim more plausible, he extends the period of time supposedly covered by the Axial Age to include the era of Jesus, and then of Muhammad. This allows him to imply that the youngest two monotheisms contributed directly to the paradigm-shift in human thought – which is more than Jaspers ever claimed. So in Habermas' recent writings, the Axis of Time extends from 800 BCE to 600 CE, a period of 1400 years which covers, perhaps, the *first half* of reliably recorded human history. One may reasonably ask if the idea of an "axis" is still meaningful here.

This transition, according to Habermas, comes about as the result of traditional society facing an economic crisis which is beyond the steering capacity of its core institution, the feudal state. What is needed is a new core institution, based on "the internal delimitation of a universalistically organized area of strategic activity (the reorganization of labour relations in the capitalist enterprise)." (HE 38) Habermas argues that the universalistic structures of consciousness developed during the Axial Age could provide this new component of the social formation. Craftsmen and merchants working in the publishing industry in various parts of Europe from the 15th century onwards[44] acted as the social movement which brought the requisite structures of consciousness out of their latency, and embodied them in capitalist firms, labour markets and commodity markets. "Only here," says Habermas, "did the universalistic structures present in all developed high cultures emerge out of their world-view latency, only here did the innovation potentials of craftsmen and in part also of mercantile elites so consolidate that universalistic principles were abstracted from the world-views and *institutionalized* in the area of social labour." (HE 38-9) In later texts, Habermas links this argument to Weber's account of the early Protestant bourgeoisie as the "carrier strata" for capitalism. (TCAI 199-200)

Habermas' idea, presumably, is that the universalistic ways of thinking dating from the Axial Age, which are connected to propositionally differentiated speech and postconventional morality, are embodied in the actions of these merchants and craftsmen. The infinite profit motive, unrestricted markets, the conception of human beings as rational utility-maximisers possessed of property rights, and so on, are all related in some way to these higher levels of communicative and moral capacity. To quote Ingram, "The uncoupling of a self-regulating market from the state enables people to coordinate their economic lives by means of standard

---

[44] Habermas in fact says the 13th century (HE 38-9, and in the original German version, "Geschichte und Evolution," *Zur Rekonstruktion des Historischen Materialismus* p 243). This is a factual error, since Gutenberg did not invent his printing press until 1440, in other words the mid-15th century.

monetary and price mechanisms without having to negotiate the exchange of commodities…

the moral bedrock underlying this norm-free strategic pursuit of self-interest is a system of

individual rights rooted in postconventional morality."[45] By displaying these ways of thinking and

acting, this "social movement" alters society's form of social integration, prompting the

adoption of a new core institution (the market) and organizational principle (the wage-labour

capital relationship). The new core institution, with its greater steering capacity, was able to

overcome the economic crisis, and in this way the traditional evolved into the modern social

formation.[46]

      This is Habermas' theory of social evolution, as I interpret it. He believes that under the

pressure of crises, social formations adopt new forms of social integration, through the

embodiment of communicative and moral capacities which lie latent in the lifeworld until they

are brought into play by social movements. This allows the social formation to form new core

institutions and organizational principles. These in turn possess greater steering capacity than

their predecessors and are thus able to overcome crises. (CES 122, 147, 160) I have not

recounted this theory for its own sake, but in order to illuminate the place of religion in

Habermas' early social theory. This will be the focus of the next few sections, which move on

from texts such as *Communication and the Evolution of Society* and *Legitimation Crisis* to *The

Theory of Communicative Action*. I will explain why Habermas associates a certain worldview

with each social formation – myth with the primitive, religion with the traditional, and a

posttraditional worldview with the modern (CES 103-4) – and go into detail about the nature of

these worldviews, which are characterised by the levels of communicative and moral

---

[45] Ingram op. cit., pp 315-6.

[46] Alan R. How interprets the trial of Louis XVI as, approximately, the moment of transition from the
traditional to the modern social formation. "In Habermasian terms," he says, the trial "signalled the
emergence of a new principle of social integration, one which differentiated culture from nature
allowing a morality of putative universal rights detached from the mythical source of the monarch's
persona to appear." Alan R. How, "Habermas, History, and Social Evolution: Moral Learning and the Trial
of Louis XVI," *Sociology* 35:1 (2001), p 191.

competence which they embody. We have seen how the centrifuge spins. It is now time to look

inside it.

1.6 Worldviews

The predominant *worldview* of each social formation correlates with the level of communicative and moral competence possessed by its form of social integration. Thus, mythical worldviews correlate with symbolically mediated interaction and preconventional morality; religious-metaphysical worldviews with an intermediate stage of communicative competence, and conventional morality; and modern worldviews with propositionally differentiated speech and postconventional morality. In sections 1.8-10 I will investigate the nature of this correlation. I will also explain why Habermas attributes certain features to each worldview. For example, in "Historical Materialism and the Development of Normative Structures," he says that each worldview has its own form of explanation and reasoning:

> Mythology permits narrative explanations with the help of exemplary stories; cosmological world views, philosophies, and higher religions already permit deductive explanations from first principles (the originary actions of myth having been transformed into the "beginnings" of argumentation, beyond which one cannot go); modern science, finally, permits nomological explanations and practical justifications, with the help of revisable theories and constructions that are monitored against experience. (CES 103-4)

So mythological worldviews explain phenomena through narrative, religious ones through deduction from first principles, and modern ones through testable laws of nature. Similarly, each worldview has its own mode of moral justification. As Habermas says in *Legitimation Crisis*:

> In the development from myth, through religion, to philosophy and ideology, the demand for discursive redemption of normative validity-claims increasingly prevails. [If we rationally reconstruct this process, we can observe] the expansion of the secular domain *vis-á-vis* the sphere of the sacred... [and the] increasing reflexivity of the mode of belief, which can be seen in the sequence: myth as immediately lived system of orientation; teachings; revealed religion; rational religion; ideology. (LC 11-2)

These developments in justification evidently track the advance from preconventional to postconventional morality which is central to Habermas' theory of social evolution, although precisely how is not yet clear. Before tackling these points, I will clarify what Habermas means by a "worldview" (*Weltbild*).

A worldview, for Habermas, is *an interpretation of a lifeworld*. It is a "cultural interpretive [system]… that reflect(s) the background knowledge of social groups and guarantee(s) an interconnection among the multiplicity of their action orientations." (TCAI 43) In most expositions of Habermas' concept of communicative action, interlocutors are simply said to draw on shared knowledge and situation definitions from their common lifeworld in the process of communication. In fact, the lifeworld as a whole is too large and all-encompassing to process; a worldview is one view of the lifeworld, a particular mapping of the terrain. It follows that several different worldviews can map the same lifeworld. A worldview's functions include making communication easier by providing interlocutors with ready-made, shared situation definitions, and thus making it possible for them to coordinate their actions and (at least under some conditions) come to rational agreements.

Habermas distinguishes his conception of worldviews from Peter Winch's Wittgensteinian conception.[47] He agrees with Winch that worldviews are totalities: "we cannot get behind them as articulations of an understanding of the world, even if they can be revised." (TCAI 58) You cannot simply step outside of your worldview, although you can (gradually and selectively) modify it. Borrowing a metaphor from Patrick Burke (TCAI 58 n 26), Habermas compares worldviews to portraits. A portrait shows a person as they appear from a certain angle; it is possible to paint several dissimilar portraits of the same person, which could all be "accurate, authentic, or adequate." "Similarly, worldviews lay down the framework of fundamental concepts within which we interpret everything that appears in the world in a specific way as something. Worldviews can no more be true or false than can portraits." (TCAI 58)

Nonetheless, Habermas differs from Winch in asserting that worldviews can be assessed with regard to truth. While worldviews as totalities may not be true or false, they do have a connection to cognitive truth, in that they are the frameworks within which we make and assess

---

[47] Habermas refers to Winch's "The Idea of a Social Science" and "Understanding a Primitive Society," in Bryan Wilson ed., *Rationality (Key Concepts in Social Science)* (Oxford, Blackwell, 1970).

true or false statements: "worldviews differ from portraits in that they in turn *make possible* utterances that admit of truth. To this extent they have a relation, albeit indirect, to truth…" (TCAI 58) Only within our worldviews do we raise the validity claim of truth, and a particular worldview may be better or worse at enabling us to do so. Habermas does not think it is possible to judge a worldview as a whole in terms of its truth or falseness, since the worldview is one of the things which allows us to thematize truth claims. At the same time, Habermas insists that truth is a universal validity claim, however it is articulated. So worldviews can be judged in terms of *whether they allow this universal validity claim to be raised*. By using this criterion of formal-pragmatic adequacy, Habermas seeks to avoid Winch's relativism. As far as he is concerned, it is possible to say that one worldview is more adequate than another. Contra Winch, "worldviews can be compared with one another not only from the quasiaesthetic and truth-indifferent standpoints of coherence, depth, economy, completeness, and the like, but also from the standpoint of *cognitive adequacy.* The adequacy of a linguistically articulated worldview is a function of the true statements that are possible in this language system." (TCAI 58-9)

Habermas uses this argument to create a criterion for evaluating the rationality of worldviews. "The rationality of worldviews is not measured in terms of logical and semantic properties but in terms of the formal-pragmatic basic concepts they place at the disposal of individuals for interpreting their world." (TCAI 45) In other words, worldviews should be ranked with regard to how developed they are in formal pragmatic terms – i.e. to what extent they enable discourse and communicative action by allowing the basic attitudes, world-concepts and validity claims to be thematized and contested. (TCAI 70) I will explain these concepts in the next section. For now, the important point is that Habermas considers worldviews which are more capable of supporting discourse and enabling communicative action to be more rational. He uses Jean Piaget's concept of *centration* to evaluate this feature of worldviews. Piaget argues that a child at the pre-operational stage of cognitive development has a "centred" understanding. She is only able to understand one aspect of a situation at a time, and is unable to switch to another

46

aspect in response to a cognitive problem.[48] In one well-known experiment, a child is shown two glasses which are the same size, filled to the same level with liquid. She is asked if the same amount of liquid is present in both glasses, and will typically be able to understand that there is. The liquid from one glass is then poured into a taller, thinner glass, so that it comes up to a higher level, and the child is again asked if both glasses contain the same amount of liquid. A child at the pre-operational level, with a centred understanding of the situation, will answer that there is now more liquid in the tall thin glass. She is fixated on one category of measurement – the height of the liquid – and is unable to switch to considering other relevant aspects of the situation, such as the difference in shape and size between the two glasses. An older child with a *decentred* understanding would not make this mistake.

Habermas uses the concept of centration to refer to the efficacy of worldviews for allowing interlocutors to reach understanding via communicative action. Worldviews are interpretations of the lifeworld; like the lifeworld, they contain the situation-definitions and interpretations passed down by previous generations. Interlocutors draw on these in discourse. They either take their interpretations and situation definitions ready-made from the lifeworld, or construct them themselves in communicative action. Interlocutors with a heavily centred worldview will have to take the former option, and those with a decentred worldview can take the latter:

> The more the worldview that furnishes the cultural stock of knowledge is decentred, the less the need for understanding is covered *in advance* by an interpreted lifeworld immune from critique, and the more this need has to be met by the interpretive accomplishments of the participants themselves... the more frequently we can expect rational action orientations. (TCAI 70)

So, individuals with a centred worldview must accept a pre-established consensus, and individuals with a decentred worldview can establish their own consensus via communicative action. The former, like the preoperational child, are fixated on one aspect of their situation,

---

[48] Jean Piaget, *Play, Dreams and Imitation in Childhood* (London: Routledge & Keegan Paul, 1962)

and are unable to consider it from other angles. The latter have a greater flexibility in how they choose to interpret the situation, and can pick out and challenge particular aspects of it.

At the communicative level, those worldviews which enable their adherents to thematize validity claims and distinguish between attitudes and worlds are more decentred, and more capable of supporting discourse and communicative action. As we will see, Habermas considers mythical worldviews to be highly centred, religious-metaphysical ones to be partly centred, and modern worldviews to be highly decentred:

> To the degree that the lifeworld of a social group is interpreted through a mythical worldview, the burden of interpretation is removed from the individual member, as well as the chance for him to bring about an agreement open to criticism… The linguistic worldview is reified as the world order and cannot be seen as an interpretive system open to criticism. Within such a system of orientation, actions cannot reach that critical zone in which communicatively achieved agreement depends upon autonomous yes/no responses to criticisable validity claims. (TCAI 71)

By contrast,

> It is distinctive of the modern understanding of the world that the cultural tradition can be exposed to testing … across its entire spectrum and in a methodical manner. Centred worldviews that do not yet allow for a radical differentiation of formal world-concepts are, at least in their core domains, immunized against dissonant experiences… What was until then "taken for granted," is transformed in the process into cultural knowledge that can be used in defining situations and exposed to tests in communicative action. (TCAII 133)

So much for Habermas' account of worldviews. It is clear that he considers the extent to which a worldview allows differentiation between the formal-pragmatic elements of language – validity claims, world concepts and basic attitudes – to be its most important feature. Before moving on to examine the particulars of the myth, religion and modernity, I must look into the details of these elements of thought and language.

1.7 Validity Claims, Basic Attitudes and World-Concepts

Habermas believes that under modern conditions speaking and acting subjects are capable of adopting three *basic attitudes*, (MCCA 137-8) which can be applied in various combinations to three *world-concepts*; an attitude applied to a world-concept constitutes a *world-relation*. These are the *objectivating*, *norm-conformative* and *expressive* attitudes. Someone adopting the objectivating attitude stands towards a state of affairs as an external observer, aiming to examine, learn about and manipulate it. This is the attitude of the scientist or engineer, and it is connected to instrumental reason. Someone adopting the norm-conformative attitude stands towards a state of affairs as a participant, who must follow rules and fulfil obligations. Failing to adhere to a norm is a matter of moral wrongness, not instrumental success or failure. As I have noted, legitimate norms are internally connected to communication in Habermas' theory. The norm-conformative attitude thus has a communicative aspect as well; it is the attitude of someone encountering and establishing a relationship with an interlocutor. Someone adopting the expressive attitude articulates and gives voice to a state of affairs. This is the attitude of the poet or artist. Aside from associating it with art and aesthetics, Habermas says little about it.[49]

Each attitude has a corresponding type of action. *Instrumental action*, in which the world is manipulated to achieve a certain end, corresponds to the objectivating attitude; *communicative action*, which aims at unforced consensus, corresponds to the norm-conformative attitude; *dramaturgical action*, which involves adopting roles and presenting a state of affairs to an audience of others, corresponds to the expressive attitude. (TCAI 90-1) *Strategic action*, in which communication is used in order to achieve an objective, is a hybrid of communicative and instrumental action and their respective attitudes.

---

[49] For an account of this underdeveloped aspect of Habermas' thought, see Pieter Duvenage, *Habermas and Aesthetics: The Limits of Communicative Reason* (Cambridge: Polity, 2003).

These attitudes give rise to three "worlds" or "formal world-concepts." (TCAI 236) The *objective* world is the external world of things ("existing states of affairs"). The *intersubjective* world is the social world which exists between communicating human beings ("legitimately regulated interpersonal relationships of a social group.") The *subjective* world is a person's internal world of feelings and drives ("experiences to which one has privileged access"). (MCCA 58) It is important to understand that Habermas is not making an ontological claim when he asserts the existence of these three worlds. (TCAI 52) As I will explain in Part 2, his approach to philosophy is postmetaphysical. What this means, among other things, is that as a postmetaphysical thinker, Habermas does not claim to have immediate access to the nature of reality. Postmetaphysical philosophical claims, he argues, must be grounded in language and the processes of communication. (PT 45) This is why Habermas insists that *attitudes are prior to worlds*. Each formal world-concept is simply a domain of reality which is delineated by an attitude. For example, the objective world is that subsection of reality which the objectivating attitude picks out.

It is the differentiated nature of modern language which makes it possible for speakers to adopt different attitudes. As I noted in section 1.4, Habermas characterises the earliest form of language, symbolically mediated interaction, as one in which "representation, expression, and behavioural expectation are one." (CES 36-7) In contrast, he considers these functions to be separated in modern language:

> The decentered understanding of the world is thus characterized by a complex structure of perspectives. It combines two things: first, perspectives that are grounded in the formal three-world reference system and *linked with the different attitudes toward the world*, and second, perspectives that are built into the speech situation itself and *linked to the communicative roles*. The grammatical correlates of these world and speaker perspectives are the three basic modes of language use on the one hand and the system of personal pronouns on the other. (MCCA 139)

The tripartite structure of attitudes and worlds is connected to the distinction between the first, second and third person which can be found in most natural languages.[50] The grammatical third person correlates to the objectivating attitude and the objective world; the second person to the norm-conformative attitude and the intersubjective world; and the first person to the expressive attitude and the subjective world. Habermas' "worlds," then, are not ontological divisions. They are the products of the ability possessed by competent language-users to switch between the three basic attitudes. By doing so they also switch between the types of action, types of argumentation and, importantly, *measures of validity* appropriate to each.

Each attitude has, along with its own world, grammatical person and so on, its own validity claim. Maeve Cooke, highlighting this last point, refers to Habermas' worlds as "validity dimensions." On this interpretation, a "world" is simply the domain of a particular validity claim. The objective world is everything which is evaluated according to the validity claim of truth, and so on. Cooke asserts that "the ability to distinguish between "worlds" is nothing more than the ability to distinguish between types of validity claims."[51] Habermas is not clear about the relations of priority between all these elements of thought and language. As I interpret him, the attitudes are fundamental, and are learned by (modern) subjects at the point of language acquisition, along with the three grammatical persons. This allows them to raise the three validity claims, delineate the three worlds, and master the three basic types of action.

---

[50] But not all. Although most Indo-European languages are confined to three grammatical persons, this practice is not universal. Finnish and Estonian have a so-called "zero person," which is used as an indefinite referent – it can mean either "I," "you" or "it/they" The closest equivalent in English is the formal "one." See Lea Laitinen, "Zero Person in Finnish: A grammatical resource for construing human reference," in Marja-Lissa Helasvuo and Lyle Campbell ed., *Grammar from the Human Perspective: Case, Space and Person in Finnish* (Amsterdam: John Benjamins, 2006), pp 209-231. In Native American languages such as Ojibwe and Navajo, the third person is divided into two – the proximate for the more salient referents of a sentence, and the obviate for the less salient referents. This obviate third person is sometimes referred to as a "fourth person." See Judith Aissen, "On the Syntax of Obviation," *Language* 73:4 (1997), pp 705-750. These examples may not derail Habermas' model, but they do suggest that the correspondence between worlds, attitudes and the structure of language is not as neat as he assumes.

[51] Maeve Cooke, *Language and Reason: A Study of Habermas' Pragmatics* (London/Cambridge, MA: MIT Press, 1997), p 11.

Habermas claims that any attitude can be applied to any of the three worlds, leading to nine *world relations*. (MCCA 137-8) The objectivating attitude, for example, need not be applied solely to the objective world; it could be applied to the subjective or intersubjective worlds. Habermas seems to envisage a two-stage process. During the first stage, the basic attitudes establish the three corresponding worlds; during the second stage, the attitudes are uncoupled and re-applied to the worlds in new combinations. This seems to be his position when he says that "facts, norms and subjective experiences have their *originary* locus in "their" corresponding worlds... and, *in the first instance*, are accessible, or identifiable, only from the perspective of an actor who takes a corresponding attitude (be it objectivating, norm-conforming, or expressive)."[52] Habermas considers six of the nine world-relations to be *rationalizable*. Each of these is the basis of a *cultural value sphere*, and these six are grouped into three *rationality complexes*. The cognitive-instrumental rationality complex comprises the cultural value spheres of science/technology and social technologies; the moral-practical rationality complex comprises the value spheres of law and morality; and the aesthetic-practical rationality complex comprises the value spheres of eroticism and art. (TCAI 238) A cultural value sphere consists of a set of ideas or a tradition, and institutions and practices in which those ideas are embodied. For example science, as it exists in modern societies, is both a set of theories and a set of institutions (research institutes, journals, universities etc.)[53] Figure 3 illustrates these relationships.

Habermas thinks that these distinct rationality complexes can exist in modern societies because modern subjects are capable of isolating and switching between attitudes and worlds. This was not always the case. The communicative capacity required is a product of social

---

[52] Jürgen Habermas, "Questions and Counterquestions," in Richard J. Bernstein ed., *Habermas and Modernity* (Cambridge: Polity, 1985), p 208.

[53] Ibid p 206.

evolution.[54] Habermas uses this schema of rationality complexes and value spheres for critical purposes. The pathologies of modernity, he argues, result from *uneven* or *selective* rationalization of world relations. One rationality complex has atrophied, or grown out of control, or the interaction between rationality complexes has broken down. (TCA1 240) As Ingram points out, this gives Habermas another opportunity to correct the pessimistic diagnoses of modernity made by Weber, Adorno and Horkheimer.[55] Modern societies are not

Worlds / Attitudes	1 Objective	2 Social	3 Subjective	1 Objective
3 Expressive	Art			
1 Objectivating	Cognitive-Instrumental Rationality Complex Science/technology	Social technology	X	
2 Norm-conformative	X	Moral-Practical Rationality Complex Law	Morality	
3 Expressive		X	Aesthetic-Practical Rationality Complex Eroticism	Art

Figure 3 – Worlds and Attitudes (based on TCAI 238, figure 11)

[54] See Ingram op. cit., pp 317-8.

[55] Ibid, pp319-20.

condemned to anomie, heteronomy or fragmentation; these effects are the products of a contingent path of uneven rationalization, which can be rectified.

Habermas considers a world relation to be rationalizable, able to support a value sphere, if it allows the "systematic production of knowledge." He thinks that only six of the nine world relations are capable of this (TCAI 237-8), as depicted in figure 3. By systematic production of knowledge, Habermas means "a production of knowledge that is differentiated according to validity claims and rendered cumulative," and whose continuity over time is guaranteed by reflexive learning processes. (TCAI 239) Moreover, Habermas thinks that the way to test whether a value sphere gives rise to such a process is to see if it has its own specialized discourse, a "typical form of argumentation specialized in accord with a universal validity claim." (TCAI 239) So we would expect to find that science, law, art and so on have their own specialized discourses, in which statements are tested against a certain validity claim, and that these discourses generate cumulative knowledge.

The three world relations which are not viable are the objectivating attitude applied to the subjective world, the norm-conformative attitude applied to the objective world, and the expressive attitude applied to the social world. Habermas does not argue that it is impossible to apply these attitudes to these worlds, or that such applications will not give rise to some social phenomena. Rather, he denies that these relations can generate systematic and cumulative knowledge in accord with a validity claim. This means that no cultural value sphere can develop from any of these relations. Habermas suggests that the objectivating attitude applied to the subjective world can be seen in empiricist psychology and utilitarian ethics[56] (TCAI 236), the

---

[56] Habermas apparently does not consider these well-established, long-running research programmes to have systematically produced knowledge differentiated according to validity claims, or to have institutionalized reflexive learning processes and developed their own specialised discourses. Certainly, as a thinker who takes a Kantian approach to morality, Habermas does not agree with the basic tenets of utilitarianism. But to ignore the historical achievements of these research programmes altogether, and dismiss them for being based on "non-rationalizable" world-relations, seems unwarranted.

norm-conformative attitude to the objective world can be seen in "mystically inspired traditions or taboos" and the anthropomorphization of animals (TCAI 236), and the expressive attitude to the social world can be seen in countercultural forms of life and, borrowing Weber's term, "Bohemianism." (TCAI 238-9)

Objections can be made to several aspects of Habermas' model. Ingram questions Habermas' assertion that the application of the expressive attitude to the social world is not rationalizable.[57] David Colclasure makes a number of criticisms of Habermas' account of art, suggesting that it is not necessarily centred on the validity claim of truthfulness, and arguing that it does not cumulatively generate knowledge.[58] McCarthy has made similar objections in the cases of art and morality, arguing that they do not progressively accumulate knowledge in the same way that science does.[59] Nonetheless, this section describes the nature of attitudes and worlds, and the relations between them, as understood by Habermas in *The Theory of Communicative Action*. As I have noted, these world-relations and rationality complexes are products of social evolution. Habermas thinks that only modern speaking and acting subjects – those with modern worldviews – have the necessary communicative capacity to differentiate between worlds, attitudes and validity claims, and establish relations between them in this way. In earlier social formations worlds, attitudes and validity claims were not so easily separable.

---

[57] See Ingram op. cit., pp 318-9.

[58] David L. Colclasure, *Habermas and Literary Rationality* (London and New York: Routledge, 2010). Both Colclasure and Ingram point out that there are harmful consequences to Habermas' decision to ascribe art, and particularly literature, to the expressive attitude, and then to argue that the expressive attitude to the social world is non-rationalizable. This seems to clip the wings of literature, and make it incapable of saying anything significant about social, rather than individual concerns (Colclasure pp 29, 35-6, Ingram pp 318-9). In fact it seems to rule out the beneficial cross-pollination between the literary and political public spheres which Habermas describes in *The Structural Transformation of the Public Sphere*.

[59] McCarthy op. cit., p 179. Habermas acknowledges some of these difficulties with regard to Weber's model of the rationalization of value-spheres – see TCAI 177.

The consequences of this lack of differentiation can be seen in Habermas' account of mythical

and religious worldviews. It is to these worldviews that I now turn.

1.8 Myth – A Syndrome of Validity

To Habermas, myth is simply the polar opposite of modernity. Modern language-users can differentiate between the three attitudes, worlds and validity claims. Mythical thinkers can do none of these things. "(M)ythical thought," says Habermas, "*does not yet permit* a categorical separation between cognitive-instrumental, moral-practical, and expressive relations to the world..." (TCAI 63)[60] The mythical conflation of attitudes, worlds and validity claims is the suspension inside the centrifuge at its most muddled. This conflation colours the types of reasoning and justification found within mythical worldviews.

Habermas takes up other theorists' accounts of mythical thought, and gives them a formal-pragmatic explanation. For example, he notes that in myths, the natural and social worlds are not cleanly separated. "From Durkheim to Lévi-Strauss, anthropologists have repeatedly pointed out the peculiar *confusion between nature and culture*. We can understand this phenomenon to begin with as the mixing of two object domains, physical nature and the sociocultural environment." (TCAI 48) To Habermas, this is evidently the result of an inability to distinguish between the objective and intersubjective world-concepts. Consider the instance in Norse mythology where all of nature – rocks, plants, fire, water – agrees not to injure the god Baldur, and then weeps when he is indeed killed.[61] Here both norm-conformative behaviour (coming to a binding agreement) and the expression of emotion (weeping to express grief at Baldur's death) are imputed to the objective natural world. This, according to Habermas, is why mythical reasoning is *analogical*. It works through "a complex of analogies in which all natural and social phenomena [are] interwoven and [can] be transformed into one another." (CES 104)

---

[60] Alexander Bertland has argued that Habermas simply treats myth as an inadequate, unreflexive version of modern thought, and does not recognize that it has an internal logic, in the form of mimesis. See Alexander Bertland, "Habermas and Vico on Mythical Thought," in Lewis Edwin Hahn ed., *Perspectives on Habermas* (Chicago: Open Court, 2000).

[61] See Barbara Leonie Picard, *Norse Gods and Heroes* (London: Oxford University Press, 1961), chap. XVIII, "The Death of Baldur."

If there is no categorical difference between the three worlds, then any phenomena within them can be connected by analogy – the rain can be equated to the sky weeping for Baldur.

Habermas says that mythological worldviews "establish an analogical nexus between man, nature and society which is represented as a totality in the basic concepts of mythological powers." (TCAII 56) Everything is connected, and the methods for attaining an end in one world can be applied anywhere. This explains the magical practices which are associated with mythical worldviews. (TCAI 47-8) "The mythical concept of powers and the magical concept of conjuring systematically impede the separation of an objectivating attitude to a world of existing states of affairs from a conformist or nonconformist attitude to a world of legitimately regulated interpersonal relations." (TCAI 49) To a mythical thinker, there is no reason why speech-acts, which can bring about changes in human society, should not have the same effect on the physical world. Habermas says that mythical thinkers confuse the symbolic causation of communication with the physical causation of objects. (TCAI 49) This renders them unable to distinguish between the normative success or failure of establishing intersubjective relations, and the physical success or failure of achieving a goal in the objective world. They apply the same monolithic conception of validity to both the social and natural worlds.

Habermas considers *narrative* to be the mythical mode of explanation. If the same reasoning applies to natural events as to human actions, it makes sense to explain the forces of nature in the same way that we would explain the activities of fellow human beings: by telling stories about their intentions and motivations. This is why "mythology permits narrative explanations with the help of exemplary stories." (CES 103) Modern thinkers explain natural phenomena in impersonal, nomological terms. Mythical thinkers personify them and tell stories about them. In Greek mythology it is Demeter's grief over the abduction of her daughter Persephone by Hades, and her periodic return to the underworld, which explains crop failures

and the seasonal death of vegetation.[62] The counterpart to the anthropomorphization of the natural world is the naturalization of the human world. Habermas thinks that by conflating the features of human society with the facticity of nature, the mythical worldview legitimises the political structures, such as they are, of the primitive social formation. This is not equivalent to political legitimation as it exists in the traditional or modern social formations, however, since primitive societies have few hierarchical structures to legitimate. As Habermas puts it, mythical worldviews "do make available a potential for narrative explanations, but they are still so tightly interwoven with the system of institutions that they explicate it rather than subsequently legitimate it." (TCAII 56) It is only later, when the primitive social formation gives way to the traditional, that mythical narratives, by now codified into religions, are used to legitimise political authority. (CES 104, 183)

Habermas sees mythical language as a manifold of interwoven validity claims. He says that "in mythical thought diverse validity claims, such as propositional truth, normative rightness, and expressive sincerity are not yet differentiated." (TCAI 50) These claims are mixed together in a "syndrome of validity." (PT 17) Mythical language is at the level of symbolically mediated interaction or the "holistic mode of speech," the lowest level of communicative competence. Habermas argues that the undifferentiated notion of validity which belongs to myth – in which "(c)oncepts of validity such as morality and truth are amalgamated with empirical ordering concepts, such as causality and health" (TCAI 50) – has the effect of immunizing mythical worldviews from criticism. His reasoning, presumably, is that this monolithic conception of validity levels out the distinction between (in reality) contingent and historically produced social relations, and the inviolable laws of nature; if the same standard of validity applies to both, it is not possible to ask whether things could be otherwise. (TCAI 51) The whole of the worldview is charged with this undifferentiated validity. "Thus a linguistically

---

[62] See G S Kirk, *The Nature of the Greek Myths* (London: Penguin, 1974), p 63, pp 249-53.

constituted worldview can be identified with the world-order itself to such an extent that it cannot be perceived *as* an interpretation of the world that is subject to error and open to criticism." (TCAI 50)

It goes without saying that Habermas considers discourse and communicative action to be impossible at the mythical level. (TCA2 159) If validity claims cannot be thematized and contested, then discourse cannot take place, and interlocutors cannot coordinate their actions by coming to unforced agreements about validity claims. In the modern social formation, communicative action and discourse generate legitimacy for institutions and confer validity on norms, form group and individual identities, and coordinate collective action.[63] Recall that in terms of communicative competence, societies with mythical worldviews are at the level of symbolically mediated interaction. (CES 36-7) In order to coordinate behaviour, foster identities and achieve legitimation, they must make use of mechanisms which are within their communicative means, such as *rituals* and *paleosymbols*. (TCAII 53-4) By the latter term, which he borrows from Durkheim, Habermas means an external focus for the attention of a social group, which unifies and coordinates the group by creating identical feelings in each member.[64] A totemic object which is used in a ritual would be a paleosymbol. As Habermas puts it, paleosymbols

> make possible a kind of intersubjectivity that is still just this side of the communicative roles of first, second and third persons, but is nevertheless beyond the threshold of sheer collective contagion by feelings… religious symbolism can be understood as the medium of a special form of symbolically mediated interaction. Ritual practice serves to bring about a communion in a communicative fashion. (TCAII 52)

Paleosymbols and rituals act as shortcuts to the intersubjective consensus which mythical thinkers do not have the communicative capacity to generate through discourse. (TCAII 188)

---

[63] Under ideal conditions. Habermas thinks that in practice, modern societies are too large and complex for discourse to play this role all the time. System integration through the non-symbolic steering media of money and power increasingly fill the gap, at the risk of colonizing the lifeworld. (TCAII 339)

[64] See Emile Durkheim, *The Elementary Forms of Religious Life* (London: Allen & Unwin, 1976).

"Paleosymbols," says Habermas, "have a meaning that is not yet modally differentiated, and like signals, they possess the power to steer behaviour." (TCAII 55)

This is Habermas' account of the mythical worldview of the primitive social formation. He sees it as highly centred, in Piaget's sense. Analogy is its form of reasoning, and narrative its form of justification. It employs ritual and paleosymbols to generate consensus, secure group and individual identities, and coordinate action. All of these features result from the undifferentiated nature of attitudes, worlds and validity claims at the level of communicative competence possessed by mythical thinkers. Myth is the equivalent of the suspension at the beginning of the centrifuge's cycle – a confused melange of entangled elements. The worldviews of later social formations, as we will see, display an increasing degree of differentiation.

1.9 Religion – The Sacred and the Profane

Finally we come to the crux of this part of the text – Habermas' account of religion. He believes that religions, along with cosmologies and metaphysical systems, (CES 103) are the worldviews of traditional social formations. "Without exception," says Habermas, "world religions developed within civilizations, that is, within the framework of societies organized around a state…" (TCAI 201) Myth gives way to religion for external and internal reasons.

Externally, the new social formation creates a demand for legitimation which the mythical worldview cannot meet. In terms of hierarchy, primitive society is relatively "flat"; but traditional society, centred as it is on the core institution of the state and the organizational principle of lordship and vassalage, creates for the first time enormous inequalities of power and wealth. This is the price of the increased steering capacity of the new social formation. Its worldview must therefore be capable of providing far more legitimation than its predecessor. Following Weber, Habermas asserts that the challenge which the religious worldview must face is that of "justifying the unequal distribution of life's goods. This *basic ethical problematic,* which bursts the bounds of myth, arises from a need for a religious explanation of suffering that is perceived as unjust." (TCAI 201)[65] Myth can reify the social world, by conflating it with the natural world (TCAII 188), but it cannot provide legitimation to the extent that is now required.

Internally, Habermas thinks that the mythical worldview becomes structurally impossible when the level of communication which supports it is surpassed. In the evolution from the primitive social formation to the traditional, the level of communicative capacity embodied in the form of social integration begins to advance from symbolically mediated interaction to propositionally differentiated speech. This is what enables society to adopt a new

---

[65] See Max Weber, "The Social Psychology of the World Religions," in H. H. Gerth and C. Wright Mills ed., *From Max Weber: Essays in Sociology* (London: Routledge & Keegan Paul, 1970). Weber thinks that the problem of undeserved suffering is the starting-point for all religions, and classifies them in terms of how they address the problem – through being theocentric or cosmocentric, affirming the world or rejecting it, etc.

organizational principle and core institution, according to Habermas. (CES 120) But the mythical worldview, as we have seen, is a result of the entanglement of attitudes, worlds and validity claims at the level of symbolically mediated interaction. Moving from this level of communication to propositionally differentiated speech means *disentangling these elements*. "(P)ropositional elements are first of all differentiated out of the holistic utterances of context-bound signal languages... this interferes with the signal-language mechanism of co-ordinating action and shakes the foundations of symbolically mediated interaction." (TCAII 30) This advance in communication means that the elements which are conflated in myth separate, and myth loses its efficacy as a worldview. Its mechanisms for achieving consensus, forming identities and coordinating action – rituals and paleosymbols – can no longer function as they did in the primitive social formation. "To the extent that participants in interaction have linguistically at their disposal an objective world to which they relate with propositions or in which they can intervene in a goal-directed manner, their actions can no longer be coordinated via signals. Only so long as the descriptive elements of meaning are fused with the expressive and imperative elements do signals have the power to steer behaviour." (TCAII 30) Presumably – although Habermas never says so explicitly – this is also the reason that the mythical mode of legitimating the social order, by conflating it with the natural order, now faces being undermined. Members of society, with ever more advanced communicative capacities, are now becoming capable of distinguish between the two worlds.

This is how religion is born, according to Habermas. We now need to know how the religious worldview is structured – what formal-pragmatic features characterise it – and how it works – what are its modes of reasoning, justification, and legitimation.

The central feature of the religious worldview, as Habermas sees it, is that it is based on a *localised* differentiation of attitudes, worlds and validity claims. The advance from symbolically mediated interaction to propositionally differentiated speech takes place in the mundane areas

of social labour. In these settings, members of the traditional social formation are increasingly able to switch between attitudes, delineate worlds and thematize validity claims:

> The further cognitive development advances, giving rise to an objectivating attitude by actors towards the world of perceptible and manipulable objects, the less communicative acts carried out with symbolic means can by themselves link together the actions of participants. To the extent that object perception and teleological action undergo development, propositional elements are differentiated out of signal language, elements that later take the explicit form of assertoric sentences and intentional sentences. (TCAII 54)

Here Habermas describes the gradual splitting-off of the objective world with its validity claim of truth, which is essential to the expansion of the social formation's steering capacity and the problem-solving abilities of individuals.

However, this formal-pragmatic differentiation has not yet gone all the way: attitudes, worlds and validity claims are still to an extent entangled. As Habermas says, "[the religious] *world order* is so conceived that ontic and normative questions are blended together." (TCAI 202) He seems to think that the validity claim of truth, with the accompanying objectivating attitude and objective world, starts to split away while the other formal-pragmatic elements remain entangled. "As we have seen speakers cannot replace the binding effects of signal languages with the public employment of *these* sentences." (TCAII 54) Members of the traditional social formation cannot, therefore, coordinate their actions, form their identities, and achieve normative consensus *exclusively* via discourse and communicative action, as do modern subjects. Habermas thinks that to accomplish these tasks, they must rely on earlier, and otherwise surpassed, ways of thinking and speaking:

> For this reason I would conjecture that there is *a split in the medium of communication corresponding to the segregation of the sacred from the profane domains of life*: religious signification, which makes possible a normative consensus and thereby provides the foundation for a ritual coordination of action, is the archaic part left over from the stage of symbolically mediated interaction after experiences from domains in which perceptible and manipulable objects are dealt with in a more and more propositionally structured manner flow into communication. (TCAII 54, emphasis added)

Attitudes, worlds and validity claims differentiate in the *profane* realm of life; but in the *sacred*, they remain entangled, just as they were in myth. The sacred is in fact a pocket of myth which survives into later eras, like a nature reserve inside a city.

The mythical condition of thought and language inside the sacred means that rituals and paleosymbols can continue to be effective in fulfilling their tasks – the tasks to which partially differentiated profane language is not yet equal. So for Habermas, the communicative basis of religion is a division of labour between the undifferentiated sacred and the differentiated profane realm, in which the former continues to produce social cohesion, secure identities and generate normative consensus in a mythical manner, while the latter deals with instrumental matters in a modern manner (i.e., by thematising the validity claim of truth.) The essence of Habermas' understanding of religion is that it is a worldview based on enclaves of undifferentiated thought and language which have survived into an age of differentiation. He says explicitly that "the ontic, normative and expressive aspects of validity (…) must remain fused together in the conception of the creator and redeemer God, of theodicy and of the event of salvation" and cannot be "separated analytically from one another" without a "radical problematization." (T 310) These religious concepts are mythical syndromes of enmeshed truth, sincerity and rightness, monolithic ideas combining objectivity, subjectivity and intersubjectivity. Andrew Edgar uses the Book of Ezekiel to illustrate this. In that text, "the report of a personal vision is at once conflated with a statement about the external world... and is taken as an endorsement of the prophet's right to speak in God's name."[66] All three dimensions of validity overlap. Again, Habermas says that "in comparison to profane action orientations, sacred ones enjoy a greater authority, even though validity spheres are less differentiated and the potential for rationality is less developed in sacred than in profane domains of action." (TCAII 189) Rationality in the discursive sense – the capacity for isolating and criticizing particular

---

[66] Andrew Edgar, *The Philosophy of Habermas* (Chesham: Acumen, 2005), p 184.

validity claims – cannot engage with the sacred, any more than it can with myth. If it was possible to thematize individual validity claims within it, it would no longer be the sacred, and would not be able to fulfil the functions of identity formation, co-ordination of action, and so on.

For Habermas religion is an intermediate stage of normative and communicative competence. It is the suspension at the mid-point of the centrifuge's cycle, as its entangled elements begin to separate.

Habermas takes the idea of the strict separation between sacred and profane from Durkheim. "In all the history of human thought," says Durkheim, "there exists no other example of two categories of things so profoundly differentiated or radically opposed to one another… the sacred and the profane have always and everywhere been conceived by the human mind as two distinct classes, as two worlds between which there is nothing in common."[67] It is ritual, according to Habermas, which maintains the segregation between the sacred and the profane. Ritual involves extremely circumscribed, formalized communication (TCAII 191-2), which disarms the critical potential of differentiated speech. "Religious discourse is closely joined to a ritual praxis that, in comparison with profane everyday praxis, is limited in the degree of its freedom of communication in a specific way. If a functionalist description is permitted, then it could be said that faith is protected against a radical problematization by its being rooted in cult." (T 310) By defanging profane speech in this way, ritual protects the sacred from being undermined through having its formal-pragmatic elements separated. So as well as securing identity and generating normative consensus, Habermas sees ritual, in the traditional social formation, as a prophylactic barrier around the sacred. He says that in the presence of ritual,

---

[67] Durkheim op. cit., pp 38-39.

"communicative actions are confined to instrumental roles, so that the influence that language

has on the validity and application of norms may be ignored." (TCAII 87)[68]

As he did with myth, Habermas explains the religious worldview's forms of

argumentation and justification through its communicative structure – the division between a

protected, undifferentiated sacred realm and an open, differentiated profane realm. The most

salient fact here is the *auratic* nature of the sacred. To members of traditional societies, the

sacred is suffused with an overwhelming but ambivalent power: it simultaneously attracts and

terrifies. Habermas takes this idea, once again, from Durkheim (TCAII 51), and links it to Walter

Benjamin's description of auratic artworks as being simultaneously present and unattainably

distant. (TCAII 49 n28)[69] This allows the sacred to ground the conventional moral code of the

traditional social formation. It is, says Habermas, the "emotionally ambivalent attitude toward

sacred objects that originally secured the recognition of moral authority." (TCAII 75) Validity

claims within the sacred are undifferentiated. The factual validity which modern language-users

would restrict to the objective world spills over into norms, and the moral validity which they

would restrict to norms infuses facts. Now in conventional morality, the established moral code

is obeyed simply because it is established. Individuals may question whether a particular act or

rule is consistent with the code, but they may not question the code itself – they may not ask

why those particular norms are the established and accepted ones. The moral code is good in

itself. The entanglement of factual and moral validity within the sacred is evidently suited to

ground this level of morality, and the auratic power of the sacred, the overwhelming normative

---

[68] At other times, Habermas seems to say that it is the *sacred* which protects *ritual*. (TCAII 60, 189, 193)
He may mean that rituals draw their normative power from the mythical background consensus of the
sacred (TCAII 69-70), just as institutionally bound speech acts draw their illocutionary force from extra-
linguistic institutions. (On this point see John Searle, *Expression and Meaning: Studies in the Theory of
Speech Acts* (Cambridge: Cambridge University Press, 1981), p 7.)

[69] See Walter Benjamin, "The Work of Art in the Age of Mechanical Reproduction," in Hannah Arendt
ed., *Illuminations* (London: Pimlico, 1999). Habermas is in fact returning to the original meaning of the
term. Benjamin notes that "the earliest art works originated in the service of a ritual – first the magical,
then the religious kind. It is significant that the existence of the work of art with reference to its aura is
never entirely separated from its ritual function." Benjamin p 217.

force which flows from its tangled validity claims, supplies conventional morality with its unquestionable character. Habermas points out, quoting Durkheim (TCAII 46), that violating a norm which is based on the sacred is not felt to be wrong because it is attached to a threat of punishment, as would be the case with preconventional morality. Violating the norm is felt to be wrong *in itself*, which is why punishment should follow. This is evidently an example of conventional moral thinking.

Habermas thinks that the dominant form of reasoning of the traditional social formation is deduction. He traces this, too, back to the formal-pragmatic structure of the religious worldview. Deduction starts from given premises, and arrives at necessary conclusions. The deductive reasoner does not question the premises themselves. The sacred is analogous to a premise, in that it is the starting-point of reasoning which is not itself questioned. Religious reasoning starts from the unquestionable sacred, which it cannot turn back and reflect upon. As Habermas says:

> The unity of rationalized worldviews that refer, in a theological vein, to the creation, or in a metaphysical vein, to the whole of what exists, is anchored in concepts like "God," "being" or "nature," in the ultimate principles or "beginnings"; while all arguments can be traced back to such beginnings, the latter are not themselves exposed to argumentative doubt. In these basic concepts, descriptive, normative, and expressive aspects are still fused; precisely in these "beginnings" there lives on something of mythical thought. (TCAI 214)

Habermas certainly thinks that deductive reasoning is more advanced than the analogical thinking of myth, but at the same time it is not as highly developed as the discursive reasoning of modernity, in which everything may be questioned. There is an evident parallel between the conventional morality and the deductive reasoning of the traditional social formation, as Habermas understands them. In both cases, an individual may employ very sophisticated reasoning – but only to test whether a given norm or assertion is consistent with an unquestioned background consensus. The background consensus itself is immune to rational interrogation. It is ultimately the division between the sacred and the profane, Habermas suggests, which grounds this way of thinking.

The contrast Habermas draws between the conventional and postconventional understandings of a norm illustrates this. (TCAII 69-70) At the conventional level, prevalent in traditional societies, the statement "It is right that (a) in (s)" derives its validity from its consistency with the ritually protected normative consensus of the sacred, which itself reaches back to mythical thinking. Actions can be criticized in terms of whether or not they are consistent with the normative consensus, but the consensus itself is beyond criticism. At the postconventional level, prevalent in modern societies, a statement such as "it is right that (a) in (s)" is understood as raising "a *contestable* validity claim vis-à-vis members of a social group – and not merely the claim that a speech act conforms with a norm, where the validity of that authorizing norm itself remains untouched." (TCAII 72) The validity of the norm is not drawn from any extra-discursive source; it is generated endogenously, in the procedure of discourse itself. In the modern social formation, as we will see, the sacred plays no role in grounding morality.

Finally, religious worldviews have their own ways of legitimating political authority and addressing questions of material inequality and suffering. As I noted at the start of this section, this is a particularly important function for Habermas, since in his schema the religious worldview emerges simultaneously with the rise of the state and the attendant increase in political and economic inequality. Habermas believes that religion legitimises political authority in a basically mythical manner, using narrative and analogy. He cites the example of the Egyptian pharaohs, who claimed to be descended from the gods Horus and Osiris, and thus to wield legitimate and unchallengeable authority. Habermas suggests that for this level of justification "narrative grounds are sufficient, viz. mythological stories." (CES 183) Only the mythical nature of the sacred makes this possible. (TCAII 89) Religious, cosmological and metaphysical worldviews, says Habermas, possess a

> more or less clearly marked, dichotomous structure that makes it possible to relate
> the sociocultural world to a world behind it. The world behind the visible world of

> this life, behind the world of appearances, represents a fundamental order; when it is possible to explain the orders of a stratified class society as homologous to that world-order, worldviews of this kind can take on ideological functions. (TCAII 189)

Political inequalities are thus reified as reflections of the divine or metaphysical order, and invested with the undifferentiated validity of the sacred. Theories such as the Divine Right of Kings or the Mandate of Heaven exemplify this kind of legitimation.[70] Habermas concedes that religions can be used to protest against inequalities, as well as legitimise them. (CES 181) In his account of the traditional social formation, however, he largely regards them as ideological, in the Marxian sense.[71] Religions give their believers an interpretation of the world into which they can fit their negative experiences. (LC 118-9) They attempt to explain apparently unjust suffering through theodicies, or with arguments to the effect that God is "testing" those who suffer.

All of this can be traced back to the monolithic conception of validity within the sacred, which endows all states of affairs with normative force. Habermas notes that there is almost no level of suffering or injustice which religious worldviews cannot process. He attributes this directly to the undifferentiated nature of the sacred and the segregation of the sacred from the profane:

> I would explain this unassailability by the systematic restrictions placed on communication… the basic religious and metaphysical concepts lay at a level of undifferentiated validity claims where the rationality potential of speech remains more tightly bound than in the profane practice of everyday life, which had not been worked through intellectually. Owing to the fusion of ontic, normative, and expressive aspects of validity, and to the cultically rooted fixation of a corresponding belief attitude, the basic concepts carried, as it were, the legitimization load of ideologically effective worldviews were immunized against objections already within the cognitive reach of everyday communication. The immunization could succeed when an institutional separation between the sacred and the profane realms of action ensured that traditional foundations were not taken up "in the wrong place"; within the domain of the sacred, communication remained *systematically restricted* due to the lack of differentiation between spheres of validity, that is *as a result of the formal conditions of possible understanding.* (TCAII 189, italics in original)

---

[70] See Francis Oakley, *Kingship* (Oxford: Blackwell, 2006).

[71] Habermas' strongest statement of this aspect of religious worldviews comes at TCAII 187-90. Here he equates the way religious worldviews veil system integration to Lukács' notion of the way capital prestructures the "forms of objectivity" through which a subject interacts with the world.

The same factors which allow religious worldviews to function as ideologies in the traditional social formation thus act to shield them from being undermined by the critical power of differentiated language.

This, then, is Habermas' original account of religion, worked out long before his postsecular turn in 2001. He sees it as a worldview characterised by a segmentation of the lifeworld into a profane region, where the differentiation of attitudes, worlds and validity claims develops apace, and a sacred region, protected by a barrier of ritual, in which these formal-pragmatic elements remain undifferentiated in a mythical fashion. This distinguishing feature of religion grounds the conventional morality and deductive style of reasoning prevalent in the traditional social formation, and provides legitimation for the political and economic inequality of this kind of society. In the last section of Part 1, I will turn to Habermas' account of social secularization: the linguistification of the sacred.

1.10 Modernity – The Linguistification of the Sacred

Habermas' theory of social evolution ends with an unambiguously secular account of modernity. As I noted in section 1.5, he believes that economic crisis (HE 38) leading to class-struggle and conflict over economic and political inequality (LC 20) eventually overwhelms the traditional social formation. The state is replaced by the market as society's core institution, and the lord/vassal relationship with the wage-labour/capital relationship as its organizational principle. The change in the form of social integration which enables this advance is an increase in communicative capacity to the level of fully propositionally differentiated speech. Members of the modern social formation can differentiate between attitudes, worlds and validity claims in the manner I described in sections 1.4 and 1.7. The religious worldview evidently cannot survive this growth of communicative capacity; the ability of modern subjects to distinguish between worlds, adopt distinct attitudes and thematize validity claims means the end of the sacred, as it existed in the traditional social formation. Borrowing once again from Durkheim, Habermas refers to this process as the *linguistification of the sacred*. (TCAII 46) I will give Habermas' account of the process, describe the features of the modern or posttraditional worldview, and finally consider the secular nature of this vision of modernity, in relation to Habermas' later postsecular turn.

Habermas describes the linguistification of the sacred as

a freeing of communicative action from sacrally protected normative contexts. The disenchantment and disempowering of the domain of the sacred takes place by way of a linguistification of the ritually secured, basic normative agreement; going along with this is a release of the rationality potential in communicative action. The aura of rapture and terror that emanates from the sacred, the *spellbinding* power of the holy, is sublimated into the *binding/bonding* force of criticisable validity claims and at the same time turned into an everyday occurrence. (TCAII 77)[72]

---

[72] Nicholas Adams is correct to say that the precise meaning Habermas gives to "linguistification" (*Versprachlichung*) is unclear. See Nicholas Adams, *Habermas and Theology* (Cambridge: Cambridge University Press, 2006), pp 81-2.

Linguistification means the breaching of the barrier between the sacred and the profane. Differentiated thought and language reach into the undifferentiated hinterland of the sacred, and disentangle the syndrome of attitudes, worlds and validity claims which sustained it.[73] This, presumably, is what Habermas means by the "radical problematization" of faith which occurs when the validity aspects of the sacred are "separated analytically from one another." (T 310) Ritual works through symbolically mediated interaction. In the face of fully discursive language, it is no longer able to play its part in segregating the sacred and the profane.

Drawing on Habermas' account of social evolution, we can divide the process of linguistification into five steps. First, profane language advances from partly (only the objectivating attitude, objective world and the validity claim of truth can be isolated) to fully (all formal-pragmatic elements are differentiated) propositionally differentiated speech. Second, a crisis threatening the traditional social formation prompts the institutionalisation of this surplus communicative capacity in the form of social integration. Third, ritual loses its efficacy as a prophylactic barrier. Fourth, propositionally differentiated speech penetrates the sacred. Finally, the entangled elements within it separate, and the sacred is no more. The crucial fact about the linguistification of the sacred, for Habermas – the crucial fact about modernity – is the "freeing up" or "unfettering" of communication. Modern subjects have the capacity to engage in communicative action, and to move to the level of discourse. Every validity claim can be thematized and challenged by a competent language-user; nothing is off-limits to discourse. This is the fundamental difference between the modern social formation and its predecessors.

---

[73] Ingram notes that Habermas' account of modernization turns on several different differentiations: of system from lifeworld, of culture, society and personality from each other, of rationalization complexes and cultural value-spheres from each other, and so on. See Ingram op. cit., pp 314-5. My focus in this part of the text is on the fundamental process of differentiation, which Habermas thinks is the basis of all these secondary processes – the separation of entangled attitudes, worlds and validity claims.

The contrast between the traditional and modern criteria for legal legitimacy, which Habermas borrows from Weber,[74] illustrates this difference. He agrees with Weber that "legal norms (as well as the creation and application of laws) that appeal to magic, sacred traditions, revelation and the like are devalued. Norms now count as mere conventions that can be considered hypothetically and enacted positively." (TCAI 162-3) In fact, Habermas sees Weber's account of the evolution of law as mapping on to his own view of the evolution of communicative and moral capacities:

> At the stage of primitive law, there is as yet no concept of an objective norm; at the stage of traditional law, norms are taken as given, as conventions that are passed on; only at the stage of modern law can norms be regarded as free enactments and judged in the light of principles that are themselves viewed as hypothetical. The rationalization of law reflects the same series of stages of preconventional, conventional and postconventional basic concepts that developmental psychology has shown to obtain in ontogenesis. (TCAI 258)

Habermas could equally have used the examples of oaths or contracts. In traditional social formations, it is the fact that oaths are sworn and contracts drawn up in the sight of God, saints or other divine figures that gives them their binding power. They borrow their normativity from the pre-established, extra-discursive sacred. Although the invocation of the sacred may continue in modern social formations (as when investitures and oaths of office involve placing a hand on the Bible), it dwindles to mere window-dressing. The validity of the oath does not come from its sacred paraphernalia, but from the discursive conditions under which it was made.

What concerns us here, then, is the transition from a worldview in which legitimacy is borrowed or inherited from a pre-established normative consensus, to one in which it is generated in discourse. To be specific, the legitimacy of a law arises from an unforced agreement

---

[74] Much of Habermas' account of the linguistification of the sacred, in fact, can be seen as a communicative "reconstruction" of Weber's theory of modernization as disenchantment. (TCAII 288) Just as he expands Marx's theory of historical materialism by adding communicative and normative elements discovered by rational reconstruction to it, Habermas expands Weber's theory of the decline of supernatural worldviews in the transition from traditional to modern society by explicating the changes in communicative competence which underlies it.

among those affected by the law to be bound by it. Discourse is the procedure for arriving at such an agreement. Habermas again follows Weber in identifying the features of modern law:

> The positivization, legalization, and formalization of law mean that the validity of law can no longer feed off the taken-for-granted authority of moral traditions but requires an autonomous foundation, that is, a foundation that is not only relative to given ends. Moral consciousness can satisfy such a requirement only at the postconventional level. It is here that there first emerges *the idea that legal norms are in principle open to criticism and in need of justification*… These posttraditional basic concepts of law and morality are first developed and systematized in modern natural law theories… what is important is that an autonomous grounding, independent of tradition, is required – in Weber's terms, that validity based on traditional consensus is replaced by validity based on rational consensus. (TCAI 260-1)

In *Between Facts and Norms*, Habermas codifies these insights about legal legitimacy in the modern social formation in his democracy principle, which states that "only those statutes may claim legitimacy that can meet with the assent of all citizens in a discursive process of legitimation that in turn has been legally constituted." (BFN 110) The important point is that in modern societies, legal and political legitimacy is no longer something extramundane or extradiscursive. It emerges within the human world, and discourse is the medium of its emergence. (TCAII 89)

Habermas draws a similar contrast between traditional moral validity, which is conventional, and its modern counterpart, which is postconventional. In traditional societies, norms borrow their validity from the sacred. In modern societies, by contrast, discourse is the *source* of normativity; it is the discursive conditions under which a norm is stated and subsequently interrogated which confer validity on it. Habermas's discourse principle codifies this insight, stating that "(o)nly those norms can claim to be valid that meet (or could meet) with the approval of all affected in their capacity *as participants in a practical discourse*." (MCCA 66, italics in original.) The change in the nature of normative validity reflects an advance in communicative competence. The reason that the sacred can confer a conventional type of validity on norms is that members of the traditional social formation are not yet able to thematize the validity claim of rightness, meaning that they cannot engage in moral discourse;

alongside this fact, the entangled nature of factual and moral validity within the sacred makes it easy to invest a social fact (such as the existing moral code) with normative force. Habermas says that "at a differentiated level of linguistic communication participants in interaction gain the freedom to say "yes" or "no" to validity claims… The binding effect of illocutionary forces comes about, ironically, through the fact that participants can say "no" to speech-act offers." (TCAII 73-4)

He also thinks that after linguistification, something of the auratic power of the sacred passes over to discursive, postconventional morality: "the *spellbinding* power of the holy, is sublimated into the *binding/bonding* force of criticisable validity claims." (TCAII 77, see also TCAII 91-2) Postconventional morality has to motivate modern subjects from within. It cannot rely on rewards and sanctions, or an unquestionable sacred ground.[75] In Habermas' formal pragmatics, the ability to thematize particular validity claims is a feature of fully differentiated language; in his discourse theory of morality, it is the fact that claims to moral rightness can be thematized and discursively challenged that gives them validity. For this reason, he sees the transition from conventional to postconventional morality as tracking the advance from partly to fully propositionally differentiated speech. The linguistification of the sacred is simply the climactic moment of this transition.

The tasks formerly fulfilled by rituals and paleosymbols – securing individual and group identities, achieving social cohesion, and so on – are fulfilled in the modern social formation (under ideal conditions) by discourse and communicative action. As Habermas puts it, "language takes over the functions of achieving understanding, coordinating action, and socializing individuals, and thus becomes the medium through which cultural reproduction, social integration and socialization take place. I traced these tendencies to the fact that the basic

---

[75] In later writings Habermas worries that postconventional morality lacks motivational power, when compared to moral codes grounded in metaphysical theories or the sacred. (See FHN 4)

religious consensus was communicatively set aflow…" (TCAII 288)[76] It is the advance from symbolically mediated interaction to propositionally differentiated speech that drives this change. The crucial difference between this aspect of the traditional and modern social formations is that language no longer merely *transmits* a pre-established normative consensus, but allows interlocutors to *bring one about*. (TCAII 107) In other words the scope for social integration, for establishing symmetrical social relations through free-flowing discourse, expands.

Of course, Habermas has an ambivalent view of modernity. The fact that modern subjects have the communicative and normative capacities for greater social integration does not mean that they will utilise them. The scope for instrumental as well as communicative reason expands in the modern era, and Habermas thinks that modernity contains a potential for oppression as well as emancipation. The modern social formation may rely on a form of social integration that works through discourse and communicative action, or it may rely on system integration through the non-symbolic steering media of money and power, which threatens to colonize the lifeworld. However it is used, Habermas thinks that this potential did not exist at all when attitudes, worlds and validity claims were entangled. The linguistification of the sacred ends in "a situation in which the reproduction of the lifeworld is no longer merely routed *through* the medium of communicative action, but is saddled *upon* the interpretive accomplishments of the actors themselves… This sort of growing autonomy can come to pass only to the extent that the constraints of material reproduction no longer hide behind the mask of a rationally impenetrable, basic, normative consensus, that is to say, behind the authority of the sacred." (TCAII 145) Contrary to the pessimistic diagnoses made by Weber, Horkheimer and Adorno, Habermas finds in the growing communicative and moral capacities of the human race a

---

[76] Habermas' assumption that ritual loses its significance for modern language-users is criticised by Adams, pp 74, 78-9.

potential for an emancipated future. Figure 4 summarizes the features of the worldviews which I have discussed in the last three sections.

We now understand Habermas' original view of religion, its nature and its decline in the modern era. This is the "secularism" which his recent writings aim to be "post." Several points have become clear during Part 1. Habermas' main philosophical projects – formal pragmatics, the discourse theory of morality, and the discourse theory of law and democracy – rest on foundations that are undeniably, unambiguously *secular*. They rely on differentiated attitudes, worlds and validity claims, the supernatant which emerges from the centrifuge at the end of its cycle. In contrast, Habermas considers religion to be an epiphenomenon of the entangled attitudes, worlds and validity claims which exist within the sacred during the traditional era. For human beings to arrive at the emancipatory possibilities of communicative

	Myth	Religion		Modernity
Reasoning/ argumentation	Analogy	Deduction		Discourse
Justification/ legitimation	Narrative	Analogy		Discourse
Consensus, identities, co-ordination	Ritual, Paleosymbols	Ritual, Paleosymbols		Communicative Action
Attitudes/ worlds /validity claims	Undifferentiated	Sacred: undifferentiated	Profane: partly differentiated	Differentiated
Morality	Pre-conventional	Conventional		Post-conventional
Communicative competence	Symbolically mediated interaction	Sacred: symbolically mediated interaction	Profane: partly propositionally differentiated speech	Propositionally differentiated speech

Figure 4 – Features of worldviews.

action, rationalised lifeworlds, and transparent, egalitarian social relations, they must leave behind the stage of communicative competence which makes religion possible; this is the implicit lesson of Habermas' social theory.

This does not mean that Habermas, even in texts written in the 1970s and 1980s, expected religious belief to vanish from society. He suggests in passing that religion in modernity becomes "a *cultural tradition* in need of being communicatively continued." (TCAII 88) What this seems to mean is that modern religion, like any other modern social phenomenon, must operate through communicative action. It can no longer rely on the extra-discursive power of the sacred, but must exert its influence through raising criticisable validity claims. Habermas says that "the forms of modern religiosity give up basic dogmatic claims. They destroy the metaphysical-religious "world beyond" and no longer dichotomously contrast the profane world to Transcendence, or the world of appearances to the reality of an underlying essence." (TCAII 195) It is not at all clear how this would work if, as Habermas has said, "ontic, normative and expressive aspects of validity" which are fused together in the basic concepts of religion cannot be "separated analytically from one another" without a "radical problematization." (T 310) His brief comments on religion after linguistification are frustratingly vague,[77] and contradict his otherwise entirely post-religious view of modernity.

The theologian Nicholas Adams says plainly that "In Habermas's account there is a clear development *from* the sacred *to* criticisable validity claims. Argumentation begins where ritual ends."[78] No informed commentator on Habermas' work would think of denying this, whatever their position on his postsecular turn. Maeve Cooke, for example, is very supportive of Habermas' postsecular project; if anything, she thinks that his model of postsecular deliberative

---

[77] See Adams op. cit., p 81: "[Habermas'] account is greatly impoverished by his reluctance to deal with the difficult case of modern religious formation."

[78] Ibid, p 78.

democracy does not go far enough.[79] Nonetheless, she asserts that "Habermas's account of communicative action is restricted to modern societies [and that] his account of communicative rationality makes most sense in connection with those modern societies in which postconventional forms of communicative action have developed."[80] The essence of Habermas' theory of social evolution, as I hope Part 1 has shown, is that as our communicative capacities grow, we become less and less reliant on pre-established or traditional norms, institutions and beliefs. Extra-discursive sources of authority, like the sacred, are devalued, and discourse itself becomes the only legitimate source of authority. Habermas' recent postsecular turn, then, is not a minor modification of his philosophical system; it is a major, radical change, which puts many of his central theories and projects into question.

The question which arises at this point is, why did Habermas change direction so drastically? What caused him to move from the secular standpoint which I have described in this part of the text, to the postsecular standpoint of his recent work? Part 2 attempts to answer this question by investigating a central pillar of Habermas' philosophical project: his theory of postmetaphysical thinking.

---

[79] See Maeve Cooke, "Violating Neutrality? Religious Validity Claims and Democratic Legitimacy," in Craig Calhoun, Eduardo Mendieta and Jonathan VanAntwerpen eds., *Habermas and Religion* (Cambridge: Polity, 2013).

[80] Cooke, *Language and Reason,* p 14. See also pp 26-7.

Part 2

Appropriating the Sacred

And Simon, seeing that the Holy Spirit was granted through the imposition of the apostles' hands, offered them money; Let me too, he said, have such powers that when I lay my hands on anyone he will receive the Holy Spirit.

Acts 8:18-19

Atheism is left defending the Maginot line of rationality while the postsecular tanks rumble on behind its lines.

Christopher Watkin, *Difficult Atheism*

2.1 Introduction – Postmetaphysical Thinking and Religion

Most analyses of Habermas' postsecular turn treat it in isolation, as a question of normative political theory. At most, some commentators point out the contrast between his recent political writings and the strongly secular cast of his early social theory, which I described in Part 1. The postsecular turn is rarely placed in the context of Habermas' broader philosophical project. This is what I will do in Part 2. Habermas has a longstanding commitment to articulating a *postmetaphysical* philosophy, which I will describe in sections 2.2 and 2.3. Postmetaphysical thinking is a deliberately modest conception of philosophy, which Habermas believes is the only legitimate way of doing philosophy in the modern era. It does not claim to access universal and necessary truths, it refrains from making metaphysical claims about the nature of reality, and it articulates no substantive conception of the good.

At first sight, this background might make Habermas' postsecular turn appear even more jarring. Why would an avowedly postmetaphysical philosopher wish to bring religious concepts into philosophical and political discourse? A nuanced understanding of

postmetaphysical thinking, however, shows that the paradigm is curiously reliant on religion and metaphysics. The middle sections of Part 2 elaborate this point. In section 2.4 I explore a central tenet of the paradigm, Habermas' belief that postmetaphysical philosophy can contribute to *moral* but not to *ethical* discourse. Due to this "ethical abstinence," modern subjects cannot rely on philosophy (Habermas' type of philosophy, at least) for ethical orientation. They must acquire ethical values from non-philosophical sources, such as art and religion. I call this process of borrowing ethical values from art and religion *ethical appropriation*, and describe it in section 2.5. The parallel process by which philosophers borrow concepts from religious and metaphysical traditions, which I call *conceptual appropriation*, is described in section 2.6. Ironically, by modestly limiting the scope of philosophy, Habermas' postmetaphysical paradigm creates the conditions for postmetaphysical thinkers to borrow concepts and values from religious and metaphysical sources. I will consider whether this combination of abstinence and dependence counts for or against the paradigm in the Conclusion.

As well as clarifying the relationship between postmetaphysical thinking and religion, these sections lay the groundwork for my assessment of Habermas' model of postsecular deliberative democracy in Part 3. The model relies on a procedure of "translating" religious contributions to the public sphere into secular language. Understanding the translation procedure is vital for any attempt to evaluate the postsecular model, but Habermas says little about how it is meant to work. The two modes of appropriation, however, are examples of how he thinks the contents of religious discourse can be taken over into secular discourse. They can be used to assess the success or failure of postsecular deliberative democracy. I will return to this point in section 3.9.

In the final sections of Part 2 I turn to the question of why Habermas, after working out a secular theory of social evolution and a postmetaphysical approach to philosophy, embraced postsecularism in 2001. I argue that this change of direction is not, as Habermas claims, primarily

a response to the burdens borne by the religious citizens of secular states, but an attempt to solve a structural problem afflicting the postmetaphysical paradigm. I call this the *anthropic problem*, and define it in section 2.7. The problem, briefly, is that postmetaphysical thinking, contrary to its own strictures, has an unstated theory of human nature. It tacitly relies on the human species having a certain *ethical self-understanding,* congruent with our particular *form of life*. Being ethically abstinent, however, postmetaphysical thinking cannot intervene in ethical discourses about the human form of life. The paradigm could not, therefore, defend its anthropic basis if it were threatened.

This might not seem like a serious problem for the paradigm, since dramatic changes to the ethical self-understanding of the species are few and far between. Human nature does not change overnight. In the years leading up to 2001, however, Habermas became concerned about two possible threats to the ethical self-understanding of the species: *hard naturalism*, which I discuss in section 2.8, and *genetic modification*, which I discuss in section 2.9. I will argue that Habermas is able to counter the first threat using postmetaphysical arguments, but that the ethical abstinence of the paradigm frustrates his attempts to argue against the latter. I conclude in section 2.10 that this failure of postmetaphysical thinking is the crucial factor behind the postsecular turn. Unable to argue against genetic modification using the resources of his postmetaphysical paradigm alone, Habermas adopts a position which allows him to make use of the powerful ethical values of religious traditions.

These points will become clearer as we proceed. Let us begin with an account of Habermas' postmetaphysical approach to philosophy.

2.2 Metaphysics and Philosophy of Consciousness

Habermas uses a highly idiosyncratic history of philosophy to ground postmetaphysical thinking. In texts written from the mid-1980s onwards, he divides the history of thought into three paradigms. The first is *metaphysics*, which he also calls "ontology" and the "philosophy of being." The era of metaphysics spans most of the history of philosophy, from the pre-Socratics to the beginning of the modern age. Habermas names Parmenides, Plato, Plotinus, Augustine, Thomas Aquinas, Spinoza, Leibniz and Hegel as metaphysical thinkers. (PT 13, 29)[81]

Metaphysics can be characterised, like all of Habermas' paradigms, by its *conception of truth* and its *method* for arriving at knowledge. For the metaphysical thinker, "(t)rue knowledge relates to that which is purely universal, immutable and necessary." (PT 13) Metaphysical knowledge is knowledge of something greater than from the facts of the everyday world. It is concerned with universal and necessary truths. Particular states of affairs in the world are subsumed under these universal truths; metaphysics privileges the One over the Many, unity over difference. (PT 120) Plato's theory of Forms is an obvious example. Metaphysics, says Habermas, allows "the human mind [to gain] an extramundane point of reference, a distancing perspective, from which the agitated in-one-another and against-one-another of concrete events and phenomena are joined together in a stable whole that is itself freed from the mutability of occurrences." (PT 118) It follows that if a thinker can grasp metaphysical truths, they have attained knowledge of the nature of the universe as a whole. Metaphysics assumes that ""the structures of being themselves are what is layed hold of in knowledge" (PT 13). Metaphysical thinkers arrive at universal and necessary truths through the method of

---

[81] Habermas seems to identify metaphysics exclusively with idealism. It is manifestly the case, however, that there have been many non-idealist schools of philosophy in the centuries between Parmenides and Hegel. Habermas admits that Aristotelian philosophy, ancient materialism and scepticism, medieval nominalism, and empiricism do not display the characteristics of metaphysics as described in this section. However, he claims that these tendencies are "antimetaphysical countermovements… [which] remain within the horizons of possible thought set by metaphysics itself." (PT 29)

contemplation.[82] (PT 32) The idea is that individuals can access the structures of reality, by mental effort alone. For Habermas this is the key feature of the metaphysical method.

Habermas classifies religion as part of the metaphysical paradigm. He includes figures such as Augustine, Thomas Aquinas, and Nicolas of Cusa in his list of metaphysical thinkers, and says that "(t)he world religions, especially the monotheistic ones and Buddhism, attained a conceptual level on a par with philosophical idealism." (PT 119) Both ways of thinking subsume the world of appearances under an extramundane unity, which can be conceived as "a world-transcendent creator-god, as the essential ground of nature, or, lastly and most abstractly, as being." (PT 30)

The metaphysical paradigm began to decline, according to Habermas, from the 17th century onward. He identifies several problems which undermined it, substantive and methodological. On the substantive level, he suggests that metaphysics' fixation on the One rendered it incapable of dealing with questions concerning the Many: "as long as the problematic of metaphysical unitary thinking remains in force, and as long as idealist modes of thought remain in use, the universal will triumph over the individual, which is banished to ineffability." (PT 157) The attempts of the late metaphysicians to reconcile the diversity of phenomena and entities with the posited unity of the universe, as seen in Hegel's *Differenzschift*, led only to paradoxes. (PT 120-1) Similarly, Habermas claims that metaphysics was unable to do justice to the uniqueness of particular entities, (PT 122) or to give a reasonable account of the material – matter is downgraded to the status of non-being, or even, in theodicies, characterised as evil. (PT 123). It is no coincidence that the metaphysical approach came under strain between 1600 and 1800, the era of the scientific revolution and the rise of the modern state. (PT 13) In

---

[82] This is an attempt at a clarification of Habermas' comments about the metaphysical method. He does not define it as clearly as the methods of the other two paradigms. "Contemplation" is to be understood in a broad sense, encompassing other methods which give the metaphysical thinker mediated access to reality, or Being, or Truth. Plato, for example, uses anamnesis and dialectic to acquire knowledge of the Forms.

both natural philosophy and jurisprudence, empirical evidence, verifiable facts and rational justification came to the fore. (PT 33) This new conception of truth, and the correct method for attaining it, injured the metaphysical conception of necessary and universal truths, arrived at through contemplation.

Habermas calls the paradigm which replaces metaphysics the *philosophy of consciousness*, "philosophy of the subject," "mentalism" or "subjectivism." It flourished between the 17th and the 19th centuries. Kant is the philosopher of consciousness *par excellence* for Habermas, (PT 124-5) but he also identifies Descartes, (PT 31) Fichte (PT 158-62) and Schelling (PT 124) as examples. As with metaphysics, Habermas identifies this paradigm by its method and its conception of truth. The metaphysical method of arriving at truth via contemplation is no longer seen as being reliable. Human beings cannot gain certain knowledge of the structure of reality through their thoughts alone; they can, however, have certain knowledge of their own thoughts*.* Any assertions made within the subjectivist paradigm must be grounded in the subject's self-consciousness, since this is only source of reliable knowledge. (PT 13) The method of philosophy of consciousness, then, is *reflection* or *introspection*. Kant employs this method in his transcendental arguments. It naturally colours philosophy of consciousness's conception of truth. The paradigm still aims at universal and necessary truths, but these truths are products of *synthetic* activity of the subject's mind. As Habermas puts it, "philosophy of consciousness (…) no longer thinks of the unity of the many as an objective whole prior to the human mind but conceives of it as a result of the synthesis executed by mind itself." (PT 124) Kant's belief that the subject constructs a unified experience of the world within the unavoidable categories of space, time, and cause and effect, would be an example. We as human beings may not be able to know about structure of reality itself, but we can know about our construction of reality, since "nature and history are given a rational structure by reason itself." (PT 34)

Habermas thinks that philosophy of consciousness has the same aims as metaphysics. Its approach is different, in that it relies on reflection rather than contemplation. Its conception of truth is substantially the same, albeit arrived at in a different way. Descartes and Kant claimed to have discovered truths which were necessary and universal; it is simply that the necessity and universality of their truths result from the way the subject's mind functions, not from the inherent structure of reality.

Habermas believes that philosophy of consciousness encounters many of the same substantive problems as metaphysics, such as ineptitude in dealing with the individual and the material. One particular flaw which he draws attention to is philosophy of consciousness' fixation on the subject/object relation. (PT 94-5) Since the paradigm builds up its picture of the objective world from the subject's self-knowledge, it sees everything in terms of the isolated subject acting on the world. As such it cannot grasp intersubjective phenomena. (PT 161)

The paradigm was eventually undermined, according to Habermas, during the 19th century. As with metaphysics, its method was brought into question by changes in society and in the sciences. The assumption that introspection could provide access to reliable, valid information became unsustainable after thinkers influenced by Hegel – Habermas mentions Feuerbach, Marx and Kierkegaard – drew attention to the historical nature of consciousness. (PT 39) The insight which these thinkers shared was that a subject's consciousness is not an eternal, ahistorical thing from which we can derive universally applicable knowledge. It is prestructured by history and social forces, and so any knowledge introspectively derived from it will be particular and localized. As Habermas puts it, "(t)he description of the entities which appear within the space of mental representations or in the stream of lived experiences remains tainted with the stigma of what is merely subjective." (PT 45) This insight was corroborated by the work of natural and social scientists, who identified other factors which can pre-structure consciousness. (PT 45) Linguists such as Humboldt and Saussure drew attention to language;

Darwin highlighted biological factors; and Freud questioned the idea that the contents of consciousness were unproblematically accessible to introspection.

These doubts eventually undermined the philosophy of consciousness. The third and final paradigm in Habermas' scheme – the one in which he locates his own work – is *postmetaphysical thinking.*

2.3 Postmetaphysical Thinking

"Postmetaphysical thinking" is as much a label for Habermas' preferred approach to philosophy as it is a paradigm or method. He asserts that there is "no alternative to postmetaphysical thinking" (PT 29) – it is simply how philosophy should be done in the modern era.[83] Habermas dates its emergence to the mid-19th century. He identifies the first generation of Hegel's critics, such as Feuerbach, Marx and (sometimes) Kierkegaard as postmetaphysical thinkers. Pragmatists such as C S Peirce and G H Mead, the later Wittgenstein, speech-act theorists such as J L Austin and John Searle, and second-generation critical theorists such as Karl-Otto Apel are also examples. Habermas considers his own work to be thoroughly postmetaphysical. Lorenz Puntel observes that "the term 'postmetaphysical,' as Habermas uses it, means only, 'beyond what Habermas calls 'metaphysics''… It cannot mean, 'beyond everything that, in the history of philosophy, has been called 'metaphysics.'"[84] In other words, Habermas defines the postmetaphysical paradigm by contrasting it with his own definitions of metaphysics and philosophy of consciousness. As with the earlier paradigms, I will characterise postmetaphysical thinking via its method and its conception of truth. I will also discuss the paradigm's conception of reason and of the role of philosophy.

The method which the paradigm uses is rational reconstruction, which I described in Part 1. As Habermas puts it, "(t)he reconstructive work of the linguist replaces a kind of introspection that cannot be readily checked on." (PT 7) Postmetaphysical thinkers do not use

---

[83] It is noticeable that Habermas does not, strictly speaking, argue for postmetaphysical thinking. He uses the (questionable) history of philosophy which I discussed in the previous section to present the paradigm as the end-point of philosophical development, but this is a selective presentation rather than an argument. It would be just as easy to present analytic philosophy, or phenomenology, or postmodernism as the culmination of the history of philosophy. Lorenz Puntel sees two "paths" to postmetaphysical thinking in Habermas' writings: the "historico-critical path," which I have just described, and the "thematic path," which consists of assumptions about the correct subject-matter for philosophy in the modern era. See Lorenz B. Puntel, "Habermas' Postmetaphysical Thinking: A Critique," (www.philosophie.uni-muenchen.de, accessed 22 August 2014), pp 6-10. Neither "path" is an argument per se.

[84] Ibid, p 6.

contemplation to access reality as a whole, or introspection to ground universal claims in their own subjectivity. They reconstruct species-wide competences by reading off the implicit knowledge human beings use in their ordinary interactions.

This method fits with the paradigm's conception of reason. Habermas says that postmetaphysical thinking has a "weak but not defeatistic [sic] concept of linguistically embodied reason." (PT 142) It does not locate reason in the structure of the world, nor the constructive powers of the subject. Reason, says Habermas, is immanent in the public, social, intersubjective procedures of human interaction. This is why he considers communicative reason to be a postmetaphysical conception of reason. It exists in the processes of communication in which we all engage; rational reconstruction can make it explicit and clarify its rules. (PT 144-6) Postmetaphysical reason is not an Archimedean point outside the human world, from which the philosopher can observe the whole of reality. It is situated in the lifeworld. Thus it is *contextual*, *historically and linguistically conditioned*, and *procedural*:

> The order of things that is found in the world itself, or that has been projected by the subject, or that has grown out of the self-formative process of spirit, no longer counts as rational; instead, what counts as rational is solving problems successfully through procedurally suitable dealings with reality. Procedural rationality can no longer guarantee an *antecedent* unity in the manifold of appearances. (PT 35)

Habermas sometimes describes the postmetaphysical conception of reason as "detranscendentalized," since it has given up the God's-eye-view of metaphysical reason. At the same time, he refers to it as "quasi-transcendental" (MCCA 203), since it retains the power to transcend particular contexts and has not become entirely immanent. Melissa Yates lists both the "detranscendentalized use of reason" and "weak transcendentalism" as characteristics of postmetaphysical thinking.[85] Postmetaphysical reason is weakly transcendental, she argues, to the extent that it uncovers (through rational reconstruction) the conditions of possibility of

---

[85] Melissa Yates, "Postmetaphysical Thinking," in Barbara Fultner ed., *Jürgen Habermas: Key Concepts* (Durham: Acumen, 2011), pp 38-9, 41-4. Yates lists five characteristics of postmetaphysical thinking. The other three are rational reconstruction, context transcending validity, and soft naturalism.

communication itself. Communicating subjects must make two assumptions: that there are shared objective and social worlds which they communicate about, and that they are engaged with other rational beings when they communicate.[86]

Postmetaphysical reason can be described as weakly transcendental in another sense. On the one hand, it employs a concept of context-transcending validity. Habermas believes that validity claims, although raised within the context of particular lifeworlds, refer to *universal* standards of objective truth, moral rightness, or subjective sincerity.[87] Postmetaphysical reason can lead us to these universal standards of validity, and is thus transcendental.[88] On the other hand, this is a "weak" transcendentalism, since it gives us *revisable* and *hypothetical* knowledge rather than absolute certainty. Yates illustrates this point by comparing Habermas' and Kant's views of the status of the truths attained by transcendental reasoning.[89] Kant, with his strongly transcendental subjectivist understanding of reason, thinks that these truths are knowable *a priori*, independently of empirical experience. They are metaphysical truths, universal and necessary. Habermas, with his weakly transcendental postmetaphysical conception of reason thinks that they have an *a posteriori s*tatus – they are reliant on empirical experience. They are hypotheses rather than absolute truths. Postmetaphysical reason rationally reconstructs empirical experience in order to attain its quasi-transcendental truths; these truths are therefore dependent on experience, and can be falsified by it.

---

[86] Ibid, pp 38-9.

[87] Universal in postmetaphysical terms, i.e. "universally valid within [the] relevant domain" of the subjective, intersubjective or objective "world." (See Yates, "Postmetaphysical Thinking," p 45).

[88] According to Charles Larmore, the fact that Habermas' conception of reason reaches for context-transcending standards of validity makes it metaphysical. See Charles Larmore, "The Foundations of Modern Democracy: Reflections on Jürgen Habermas," *European Journal of Philosophy* 3:1 (1995), pp 59, 62.

[89] Yates, "Postmetaphysical Thinking," pp 41-2. Here Yates is drawing on Thomas McCarthy's analysis in *Ideals and Illusions* (Cambridge, MA: MIT Press, 1993).

This brings us to Habermas' postmetaphysical conception of truth. Postmetaphysical reason does not deal in universal or necessary truths – it says silent about the nature of reality or the structure of the objective world. Its claims are limited to the intersubjective world, the lifeworld of human society. They are internal to the community of communicating human beings, while still being "universal" in the weakly transcendental sense described above. The truths of postmetaphysical thinking are therefore *quasi-transcendental, situated* and *falsifiable*. In the postmetaphysical paradigm, "cognitive accomplishments can now prove themselves only through procedural rationality, ultimately through the procedure of argumentation." (PT 38) Postmetaphysical claims have the status of hypotheses – valid until disproven, and valid because they are open to being disproven. (T 304-5) They do not claim to be objective, like metaphysical truths, but to possess "(t)he intersubjective validity of observations that can be ascertained through experimental practice and thus through a regulated transformation of perceptions into data." (PT 45) The situated aspect of this conception of truth is linked to the contextual nature of postmetaphysical reason; its falsifiability is linked to its reliance on empirical experience via the reconstructive method.[90]

The role of postmetaphysical philosophy, according to Habermas, is to act as a mediator between the lifeworld and the "expert cultures" – activities with specialized discourses such as the sciences, criticism, law and so on, based on the rationalizable value spheres. (PT 38-9) He

---

[90] Is it possible for a truth to be simultaneously transcendental (even "weakly" so) and falsifiable by empirical experience? Habermas does address this question (see BNR 41), but doubts remain about the coherence of his postmetaphysical conception of truth. Karl-Otto Apel considers Habermas' standard of falsifiability to be unworkable, arguing that "the unavoidable presuppositions of argumentation (which cannot be denied without committing a *performative self-contradiction*) cannot be fallible and subject to empirical tests, because in case of falsification they would simultaneously be presupposed in their transcendental function." See Karl-Otto Apel, "Regarding the Relationship of Morality, Law and Democracy: On Habermas' *Philosophy of Law* (1992) from a transcendental-pragmatic point of view," in Mitchell Aboulafia, Myra Bookman and Catherine Kemp eds., *Habermas and Pragmatism* (London and New York: Routledge, 2002), p 19. Puntel makes an extended critique of this aspect of postmetaphysical thinking, arguing that none of Habermas' attempts to outline a theory of truth (at least since abandoning the consensus theory described in "Wahrheitstheorien") have been successful. See Puntel, "Habermas' Postmetaphysical Thinking," passim.

compares this to the way art criticism mediates between art and ordinary life. In this way, postmetaphysical philosophy helps to keep expert cultures in contact with the common sense of the lifeworld, while simultaneously "reforming" (rationalizing without colonizing) the latter. (T 310)[91] As Habermas puts it:

> What remains for philosophy, and what is within its capabilities, is to mediate interpretively between expert knowledge and an everyday practice in need of orientation. What remains for philosophy is an illuminating furtherance of lifeworld processes of achieving self-understanding, processes that are related to totality. For the lifeworld must be defenced against extreme alienation at the hands of the objectivating, the moralizing, *and* aestheticizing interventions of expert cultures. (PT 17-8)

With its limited tools, postmetaphysical philosophy must protect and reform the lifeworld.

Habermas' paradigm is notable for what it does *not* do. It cedes its jurisdiction over factual questions to the sciences, (PT 14, 16) and makes no claims about the objective world or the nature of reality. Many of the traditional fields of philosophy are closed to the postmetaphysical thinking. The paradigm cannot endorse positions such as idealism or materialism, theism or atheism, and so on. It cannot argue against them, either – it simply abstains from making any metaphysical assertions. It cannot construct philosophical anthropologies or theories of the good.[92] In normative terms, postmetaphysical philosophy can clarify and participate in *moral* discourses, but must abstain from what Habermas calls *ethical* discourses – a significant prohibition, which I will discuss in the next section.

Habermas considers postmetaphysical thinking to be an advance on the previous paradigms. Traditional metaphysics privileges the One over the Many; reactions against

---

[91] Habermas' essay "Philosophy as Stand-In and Interpreter," MCCA pp 1-20, deals directly with this function of postmetaphysical philosophy.

[92] Of course, critics of Habermas would argue that he does not in practice adhere to his own standards – that he has a tacit metaphysics, account of human nature and theory of the good. As we have seen, Larmore and Apel make arguments of this kind. I will put these arguments aside for now, and return to them in § 2.7. Whatever the truth of Larmore or Apel's claims, Habermas certainly *thinks* that he is working within the postmetaphysical paradigm. He takes its constraints seriously, as we will see.

metaphysics, such as Adorno's negative dialectics (PT 37, 144-5) and Richard Rorty's contextualism (PT 135-6) simply invert the problem, and privilege the Many over the One.[93] Postmetaphysical thinking, according to Habermas, is able to hold the balance between the One and the Many. This is exemplified by the "quasi-transcendental" status of validity claims, which are raised within particular, local lifeworlds (the Many) but reach for universal standards of validity (the One).

Habermas offers George Herbert Mead's account of subjectivity as an example of the postmetaphysical paradigm succeeding where its predecessors failed. He suggests that philosophy of consciousness, despite its focus on the individual subject, could not adequately explain the origin of subjectivity. Subjectivist thinkers either took Fichte's route, and saw the subject as posited by itself (PT 161), or Kierkegaard's route, and conceived of it as being posited by an absolute other, namely God. (PT 24-5)[94] Habermas says that he has no objection to contemporary metaphysical thinkers like Dieter Henrich taking Kierkegaard's religious route, and understanding subjectivity as something posited from outside. "Even the rhetorical force of religious speech retains its right, as long as we have not found convincing language for the experiences and innovations conserved in it. One will nonetheless be allowed to note that the original Fichtean problem is rendered pointless by a change of paradigm." (PT 25) Mead's postmetaphysical approach sidesteps from the dichotomy of self-positing and divine positing,

---

[93] Habermas thinks that modern attempts to revive traditional metaphysics repeat the mistake of privileging the One over the Many. He also associates such attempts, by thinkers such as Dieter Henrich (PT 10-27) and Odo Marquand, (PT 133 n20) with anti-modernism and political reaction. Henrich replied to this charge in "What is metaphysics – What is modernity? Twelve theses against Jürgen Habermas," in Peter Dews ed., *Habermas: A Critical Reader* (Oxford/Malden, MA: Blackwell, 1999). Henrich criticises the postmetaphysical account of subjectivity – that the individual subject emerges from intersubjective processes of communication – and defends the subjectivist version. He also identifies problems with Habermas' soft naturalism, which I mention in § 2.8. The essay "Metaphysics after Kant" (PT 10-27) is Habermas' response to Henrich.

[94] Here Habermas classifies Kierkegaard as a subjectivist, rather than an early postmetaphysical thinker. He is inconsistent on this point, referring to Kierkegaard in a later text ("Are There Postmetaphysical Answers to the Question: What is the 'Good Life'?", FHN 1-15) as a postmetaphysical thinker. As we will see in § 2.9, Kierkegaard's location in the schema has consequences for Habermas.

by depicting subjectivity as an internalization of intersubjectivity. He argues that the self emerges from the internalization of the roles and perspectives used in communication. Individuals learn to view themselves from the perspective of the "generalized other" of the group(s) to which they belong, and thus regard their own individuality in a reflexive manner. As Mead puts it, "I know of no way in which intelligence or mind could arise or could have arisen, other than through the internalization by the individual of social processes of experience and behaviour(.)"[95] Habermas takes this to mean that within the postmetaphysical paradigm, the antinomy of Fichte's and Kierkegaard's positions has been dissolved. We do not have to choose between self-positing and divine positing – *we posit each other*, since our subjectivity arises through the internalization intersubjective communicative roles. (PT 25-6) Figure 5 summarizes my account of Habermas' three paradigms.

The postmetaphysical paradigm sets the tone for Habermas' philosophical project. He regards his major theories – of communication, morality, and democracy – as examples of postmetaphysical philosophy. His approach to perennial philosophical issues, such as the ethics/morality distinction, is shaped by his commitment to the paradigm. It is to this issue which I now turn.

Figure 5. Summary of the three paradigms.

Characteristics Paradigm	Method	Conception of Truth	Thinkers	Examples
Metaphysics	Contemplation	Universal, necessary, to be found in the world	Parmenides, Plato, Plotinus, Leibniz, Spinoza, Hegel	Plato's theory of Forms
Philosophy of Consciousness	Introspection/ reflection	Universal, necessary, produced by the subject	Descartes, Kant, Fichte, Schelling	Kant's transcendental arguments
Postmetaphysical Thinking	Rational reconstruction	Quasi-transcendental, falsifiable, situated	Peirce, Mead, Wittgenstein, Austin, Searle, Apel, Habermas	Mead's theory of subjectivity

---

[95] George Herbert Mead, *Mind, Self and Society* (Chicago: Chicago University Press, 1967), pp 191-2.

2.4 Morality and Ethics

Habermas is not the only philosopher to differentiate between morality and ethics. John Rawls, for example, draws a very similar distinction between "the right" and "the good."[96] The distinction is not even a product of the postmetaphysical paradigm. Plenty of metaphysical and subjectivist thinkers have made use of similar distinctions – consider Hegel's concepts of *Moralität* and *Sittlichkeit*.[97] It is not the distinction itself but the *role Habermas believes philosophy should play in it* which can be attributed to postmetaphysical thinking. He believes that postmetaphysical philosophy can participate in moral discourses, but not in ethical ones. I will first set out Habermas' definitions of both kinds of discourse, then explain why postmetaphysical philosophy must be ethically abstinent in this way.

Moral and ethical discourse are both varieties of *practical discourse*.[98] The former is universalist and concerns the *right*, while the latter is limited to a certain individual or community and concerns the *good*. (MCCA 178) Moral discourse concerns *norms* or *principles*, ethical discourse concerns *values*. In *Between Facts and Norms,* Habermas distinguishes between them in four ways.

First, norms are deontological, in the sense that they "obligate their addressees equally and without exception to satisfy generalized behavioural expectations," whereas values are teleological, in the sense that they are "intersubjectively shared preferences" which "express

---

[96] John Rawls, "The Priority of Right and Ideas of the Good," *Philosophy and Public Affairs* 17:4 (Autumn 1988). Like Habermas, Rawls thinks that the right takes priority over the good.

[97] GWF Hegel, *Outlines of the Philosophy of Right* (Oxford: Oxford University Press, 2008), §141 (pp 151-3). There are differences between the two pairs of concepts. For example, Hegel thinks that ethical life gives abstract morality its substantive content. Habermas would not endorse this interpretation.

[98] Habermas also discusses a third variety of practical discourse: *pragmatic discourse*, which deals with questions of the most effective way to achieve a given end. (JA 2-3)

the preferability of goods that, in specific collectivities, are considered worth striving for and can be acquired or realized through goal-directed action." (BFN 255)[99]

Second, norms are binary – they contain "a binary validity claim" which their addressees can respond to with either a yes or a no, and are therefore "either valid or invalid." Values are graduated – they "set down preference relations telling us that certain goods are more attractive than others; hence we can assent to evaluative sentences to a greater or lesser degree." (BFN 255)

Third, valid norms are absolutely binding. "The 'oughtness' of binding norms has the absolute sense of an unconditioned and universal obligation: what one 'ought to do' claims to be equally good for all." Values are only relatively binding. They have "the relative sense of an estimation of goods that has become established or been adopted in cultures or forms of life: serious value choices or higher-order preferences tell us what is good for us (or for me) overall and in the long run." (BFN 255)

Fourth, a group of norms which "claim validity for the same circle of addressees" must be coherent with each other. They must "fit together into a coherent complex, that is, form a system." This is not the case for values. They can compete with each other, and "to the extent that they find intersubjective recognition within a culture or form of life, they form flexible configurations filled with tension." (BFN 255)

Since their objects are so different, it follows that moral and ethical discourse serve different purposes and work in different ways. "In the light of norms," says Habermas,

> I can decide what action is commanded; within the horizon of values, which behaviour is recommended… if we start with a system of valid norms, that action is 'right' that is equally good *for all*; in reference to a typical value constellation for our culture or form of life, on the other hand, that behaviour is 'best' that, on the whole and in the long run, is good *for us*. (BFN 256)

---

[99] Clearly, this is a deontological rather than a consequentialist view of morality, suited to Habermas' Kantian-pragmatist approach.

This passage brings out the most important distinction between morality and ethics, from Habermas' point of view. Moral discourse aims at clarifying what is right, for *all* moral subjects in *all* circumstances; ethical discourse aims at clarifying what is good, for a *particular* individual or community. William Rehg explains this crucial difference by saying that the two discourses answer different questions. He suggests that "Who am I? (or, Who are we?) and Who do I/we want to be?" would be ethical questions. Rehg adds that these questions are closely tied to an individual's or group's identity, history or traditions – they are localized questions which do not have universally valid answers. As examples of moral questions, Rehg offers "What behaviour may we legitimately expect of one another? or, What norm or obligation can each person accept on the basis of good reasons?" The ultimate moral question of the modern era, he suggests, is "What rules governing our lives together are equally good for all?" Moral and ethical discourse not only ask different questions, they ask them from different perspectives. The perspective of ethical discourse, says Rehg, is egocentric or ethnocentric – it stays within the horizon of an individual life or the history of a community. Moral discourse, in contrast, has a universalist perspective. Since it aims to find norms which are equally acceptable for all, it transcends the bounds of individual lives and particular traditions.[100]

Habermas says a great deal about the workings of moral discourse. His discourse theory of morality[101], with its (D) and (U) principles, aims to clarify the conditions under which moral discourse takes place, so as to give rise to valid norms. I discussed this theory in Part 1. He says relatively little, in contrast, about the details of ethical discourse. In the *Theory of*

---

[100] William Rehg, *Insight and Solidarity: The Discourse Ethics of Jürgen Habermas* (Berkley: University of California Press, 1997), pp 94-5. Rehg is not using "ethnocentric" in a pejorative sense here – he is simply saying that ethical discourses inevitably centre on particular communities.

[101] "Discourse ethics," by Habermas' own admission, is a misnomer, since the theory concerns morality rather than ethics. See his "Preface" to *Justification and) Application.* (JA vii) Habermas coined the term before fully developing his account of the ethical/moral distinction, with the essay "On the Pragmatic, the Ethical, and the Moral Employments of Practical Reason" (JA 1-17), originally delivered as a lecture in 1988.

Communicative Action, which he wrote before fully developing his account of the distinction, Habermas does not consider the discussion of ethical questions to be rational discourse at all, comparing it to aesthetic criticism. (TCAI 20) Later, he conceives of it as a process of reflection, by individuals and communities, on their identity and traditions.

Habermas in fact describes two varieties of ethical discourse.[102] *Ethical-existential* discourse concerns an individual's life-project, and aims at "hermeneutic clarification of an individual's self-understanding and clinical questions of a happy or not-failed life." (JA 8-9) When I ask what kind of person I am, or what kind of life I should lead, I am asking ethical-existential questions. *Ethical-political* discourse concerns the nature and goals of a community, and aims at "the clarification of a collective identity that must leave room for the pursuit of diverse individual life projects." (JA 16) When we ask ethical-existential questions we aim to clarify who we are and who we want to be, not as individuals, but as "members of a family, as inhabitants of a region, or as citizens of a state." (JA 23) Habermas sometimes says that ethical-existential discourse takes place in the first person singular, ethical-political discourse in the first person plural. (JA 15)

Ethical discourse does not work in the same way as moral discourse. Since its objects are values or "strong evaluations" (JA 4), the (D) and (U) principles, designed to test the validity of action norms, do not apply. The standard by which values are judged in ethical discourse is that of "an authentic way of life," rather than universal acceptability. (JA 23) Habermas has never devised a "discourse theory of ethics" of comparable complexity to his discourse theory of morality, to describe the ins and outs of ethical discourse. His remarks on the topic suggest that ethical discourse involves overcoming illusions about one's identity and values. (JA 4) Persons engaging in ethical discourse must distance themselves from their own experiences, becoming reflexive about them in order to "reconstruct" and "critically appropriate" their life-history or

---

[102] Habermas' later addition of a third variety, *species-ethical discourse*, complicates the picture, as I explain in § 2.7.

the history of their community, a process which is only possible within the horizon of their shared lifeworld. The goal of this kind of discourse is to generate "evaluative statements about what is good for a particular person," or community. (JA 11-2) Habermas says that in this process "clinical" or "therapeutic" knowledge is more useful than (postmetaphysical) philosophical knowledge.

The relative paucity of Habermas' account of ethical discourse illustrates an important point: he believes that *postmetaphysical philosophy cannot contribute to ethical discourses, only to moral ones*. It "is no longer in possession of an affirmative theory of the good life," (PT 50) and thus cannot participate in deliberations about the good for individuals and communities. Its role is limited to facilitating moral discourse. "What moral *theory* can do and should be trusted to do," says Habermas, "is to clarify the universal core of our moral intuitions and thereby to refute value scepticism. What it cannot do is to make any kind of substantive contribution." (MCCA 211) Why does he believe this?

Habermas' discourse theory of morality, as we have seen, aims to be thoroughly postmetaphysical. It claims to rationally reconstruct the normativity which is built into communicative action, the fundamental form of human action. (MCCA 197-8) The theory fashions this implicit element of the situated, historically conditioned lifeworld into a set of principles which have weakly-transcendental validity. Habermas evidently thinks that morality can be approached from a postmetaphysical direction. He certainly believes that postmetaphysical philosophy should articulate and clarify the procedures of moral discourse.

Ethics, by contrast, seems to Habermas to have a tinge of metaphysics. He assumes, according to Rehg, that the "intersubjective binding force" of conceptions of the good is rooted in "religious or metaphysical world view(s)."[103] It is certainly true that many traditional visions of the good are derived from metaphysical and religious beliefs, from Aristotle to Aquinas to

---

[103] Rehg, p 97.

Hegel. Ethical discourses can be more or less reflexive. (JA 11-2) But, for Habermas, even the most reflexive ethical discourse has more in common with the surpassed paradigms of philosophy than with postmetaphysical thinking. To affirm that a certain way of life or a certain form of society is ideal and must be striven for, is to make a kind of metaphysical claim about human nature, or the telos of history, or the good. This seems to be Habermas' assumption. As a result, his general precept to eschew "metaphysics" (in his sense of the word) disqualifies postmetaphysical philosophy from addressing ethical questions. This means that the paradigm must ignore many of the issues with which philosophy has traditionally dealt. Habermas does not consider this to be a failure on the part of postmetaphysical thinking. "Any ethics that is at once deontological, cognitivist, formalist, and universalist ends up with a relatively narrow conception of morality that is uncompromisingly abstract." (MCCA 210) It is a price worth paying for the advance from metaphysical to postmetaphysical thinking.

There are other reasons for Habermas to adopt his position on postmetaphysical thinking and ethics. Ethical values are ill-suited to the methodology of the postmetaphysical paradigm. Rational reconstruction can only be applied to species-wide capacities, general cognitive structures, and so on. Ethical values are bound up with particular cultures and traditions, with specific lifeworlds, and thus are not amenable to rational reconstruction. As such, they cannot be made into objects of postmetaphysical philosophical inquiry. (MCCA 121)

In terms of Habermas' theory of social evolution, ethics seems to belong to the conventional level of the traditional social formation, morality to the postconventional level of the modern social formation. Ethical discourse is centred on a particular individual or community; conventional morality takes the values of a particular community as its touchstone. The traditional ethical consensus is grounded in metaphysical or religious belief-systems, such as the Sacred. (TCAII 46, 54-5) At this level, questions of the good for an individual or a community cannot be disentangled from questions of what is equally good for all, or right.

(MCCA 164) Morality and ethics, norms and values, are not yet separable. The advance to the postconventional level makes it possible to separate morality from ethics, to achieve a "moral point of view," from which normative questions can be judged impartially and universal answers can be found. (MCCA 198) "This 'moral point of view'," says Habermas, "projects a sharp but narrow circle of light which throws into relief, against the mass of all evaluative questions, those action conflicts that can be *resolved* in relation to a universalizable interest: the questions of justice."[104] Habermas thinks that the emergence of a separate moral discourse, explicable in postmetaphysical terms, is one of the fundamental achievements of modernity. In order for philosophy to address ethical questions, then, it would have to regress to the level of metaphysics.[105]

The upshot of Habermas' interpretation of the morality/ethics distinction, in the light of the postmetaphysical paradigm, is his injunction that postmetaphysical philosophy should stay away from ethical discourses. It cannot provide any inputs for them, any ethical values or conceptions of the good. What it can do is clarify the conditions of moral discourse. This might seem like a modest and sobering, but unproblematic, understanding of the role of contemporary philosophy. As we will see in the last sections of Part 2, however, Habermas' insistence on the ethical abstinence of postmetaphysical philosophy has serious implications for the paradigm. It may be that ethical discourse is more urgent, and harder to avoid, than Habermas imagines. For now, I will turn to the question of the sources of ethical values which are available in the modern era. Modern subjects still need to engage in these discourses, as individuals and as members of communities. They must get their ethical values from somewhere – if not from philosophy, then

---

[104] From *Die nacholende Revolution*, p 118, quoted in Rehg, p 97.

[105] The disadvantage of this transition, as I noted in § 1.10, is that postconventional moral norms are at risk of losing the immediate motivating force of ethical values. (MCCA 108-9, 178-9)

from some metaphysical source which is still present in modern societies. This is what I will investigate in the following section.

2.5 Ethical Appropriation

Habermas thinks that there are two sources for the ethical values which postmetaphysical philosophy can no longer provide: *art* and *religion*. He does not develop a clear theory of the ethical function of art. He says that both religion and "art that has become autonomous" can help postmetaphysical thinkers deal with "explosive experiences of the extraordinary,"[106] (PT 51) and suggesting that confessional literature in the tradition of Rousseau should be seen as a secularized form of prayer, in which "the prayer was deflated into a public conversation." (PT 165-7) More importantly, art and religion can supply ethical values to modern subjects. Habermas sees artworks as sources for visions of the good. He refers to Charles Taylor's belief in literature's ability to describe the good, with its "world-disclosing power of… evaluative language." (JA 74) Taylor mentions Rilke, Pound, Lawrence and Mann in particular. Although he has some reservations about the power of literature, Habermas agrees with Taylor that "the rhetorically moving, exemplary representations of the novelist" trump postmetaphysical philosophy in this respect. (JA 75) The idea seems to be that artworks can describe processes of ethical-existential struggle and depict the good, positively or negatively,[107] tasks that postmetaphysical philosophy cannot accomplish. Modern subjects then appropriate these depictions and use them as inputs in their ethical discourses.

The same seems to hold for religion. For the purposes of ethical discourse under postmetaphysical conditions, religious traditions are funds of ethical values and visions of the good. Habermas speaks of "religious and classical-philosophical ethics that take ethical life as

---

[106] Puntel objects to William Mark Hohengarten's translation of this passage, and suggests that "auβeralltäglichen" should be rendered literally as "extra-everyday," rather than "extraordinary." See Puntel, "Habermas' Postmetaphysical Thinking," p 7.

[107] By this I mean that artworks may present visions of the good positively, i.e. fictionally depict unalienated lives or communities, or they may depict alienated lives and communities in such a way as to indicate what the
good could be. In practice, works art and literature are more likely to take the negative path than the positive, outside of the historical genre of Utopias.

their theme. They too understand and justify the moral not in its own terms but within the horizon of a larger soteriological or cosmological whole." (MCCA 164-5) Of course, these inputs to ethical discourse are contestable, and are no longer imbued with the overwhelming power of the Sacred which they possess in traditional societies.

The constraints of Habermas' postmetaphysical paradigm, then, lead to modern subjects appropriating contents from religion and art for use in their ethical discourses. I will call this process *ethical appropriation*. It has been noticed by commentators on Habermas' work, but rarely investigated in detail. Maeve Cooke is one of the few who have tackled it, and I will take her account of ethical appropriation as the basis of my own. She argues that all representations of the "transcendent ethical object," or the good, have the status of fictions. Habermas' original conception of the Ideal Speech Situation is one such fiction. It is a "representation of a transcendent object that conceals its own impossibility, without compromising its motivating power."[108] Cooke suggests that it is the persuasive capacity of these ethical fictions, rather than their realizability, which makes them powerful. They can be evaluated in terms of how persuasively they "disclose" the transcendent ethical object and how effectively they provide ethical orientations.[109] Cooke agrees with Habermas that "under conditions of modernity, art, music, literature and religion have become privileged sites for disclosure," although she criticizes him for restricting ethical disclosure to these areas alone.[110]

---

[108] Maeve Cooke, *Re-Presenting the Good Society* (London and Cambridge, Mass.: MIT Press, 2006), p 118. Cooke takes the term "transcendent ethical object" from Ernesto Laclau, and bases her account of fictions on the work of Hans Vaihinger.

[109] Ibid, pp 120-1. Artworks such as paintings and pieces of music may be excellent at disclosure, according to Cooke, but since they do so in a non-verbal way, the ethical orientations they disclose are not open to being thematised and challenged. Thus they may fall short of the standards of discursive, postmetaphysical reason. Traditional religious values, on the other hand, may provide very powerful orientations, but those orientations may only function at the conventional level of morality – as such, they also fail to meet the standards of postmetaphysical thinking as well. Cooke suggests that a successful representation will have to be successful in both dimensions.

[110] Ibid, p 123, see also n 63.

She says that art and religion disclose ethical values "by means of linguistic, visual, aural, and other kinds of techniques (using, for example, shock effects, unexpected juxtapositions, and innovative metaphors) and by means of exercises such as prayer and meditation."[111] In her 2011 paper "Translating Truth,"[112] Cooke argues that ethically appropriated religious ideas possess "exemplary force" – like good examples, they allow us to point at a certain form of life or vision of the good and say, "*that* is the best example of *that* thing – *that* is how things should be." Borrowing from Alessandro Ferrara, Cooke describes this as the force of what *is* and what s*hould be*, united in the force of what *is as it should be*. This can serve as an account of how appropriated ideas provide ethical input – they act as examples.

Cooke offers two examples of ethical appropriation, which work in different ways. Departing from her account, I will classify them as different types of ethical appropriation. One example is of a real incident, the resignation in 2010 of Margot Käβmann, the Protestant bishop of Hanover, immediately after being arrested for drunk driving. Cooke notes that bishop Käβmann's action was described as "exemplary" by *Der Spiegel*. She suggests that those hearing about the incident will have little difficulty in appreciating its exemplary force and deriving an ethical value from it. This is because

> it is taken from a context with which we are all likely to be familiar: a contemporary liberal democratic European state in which there are laws about drinking and driving... If we feel the full force of the example in the Käßmann case, it is because we already inhabit the complex interpretative framework in which her act is embedded and are familiar with the linguistic expressions and modes of articulation whose interplay mediates the exemplarity of her act. More specifically, it is because

---

[111] Ibid, p 123.

[112] Maeve Cooke, "Translating Truth." *Philosophy and Social Criticism* 37.4 (2011) 379-491. As the title implies, Cooke refers to the process as "translation" rather than "ethical appropriation." This is unobjectionable in the sense that Habermas himself sometimes calls it translation, but it may cause some confusion in the context of my argument. I draw a sharp distinction between the processes of borrowing from religion under postmetaphysical conditions, described in this section and the next, and the process of producing secular versions of religious inputs to the public sphere, described in Part 3. (See § 3.9 for my argument that the processes are different.) I refer to the former as "appropriation" and the latter as "translation." Cooke does not draw such a precise distinction. The benefit of drawing this distinction is that it allows me to investigate whether (post-2001) postsecular translation and (pre-2001) ethical appropriation do, in fact, work the same way.

we have internalized key elements of this interpretative framework, such as the Kantian idea that human beings are morally autonomous agents who freely subject themselves to the moral laws that bind them.[113]

So in this case, the ethical appropriation of the exemplary act is made easy by the fact that the appropriators understand the context in which it took place; readers of *Der Spiegel* share a historical and social context with bishop Käβmann. As a result, it would not be difficult for them to appropriate ethical values from her act – "take personal responsibility for your actions," for example, or "do not shirk negative consequences when they are deserved" – and use these in their ethical discourses. I will refer to this as *type one ethical appropriation*.

Cooke's other example is very different. She discusses an incident in Orhan Pamuk's novel *My Name is Red*, which concerns manuscript illustrators in 16th century Istanbul – a cultural context remote from that of modern day ethical appropriators. Cooke focuses on an incident in which the character of Master Osman, inspired by his predecessor Great Master Bihzad, blinds himself with a gilded needle in order to be a better illustrator. Osman, who shares much of Bihzad's cultural context, finds the latter's act exemplary. He could be said to have ethically appropriated it. Modern readers, in contrast, are likely to find the incident strange and horrifying, rather than inspiring. They will find it much more difficult to make an ethical appropriation of Great Master Bihzad's act than bishop Käβmann's. "The difference," says Cooke,

> clearly has something to do with context… To feel the full force of the example in the Bihzad case, we would have to inhabit a complex interpretative framework in which the purpose of illumination is to evoke and honour the divine, in which the hand of the master illuminator is merely a vehicle for the divine, in which blindness is a gift of God bestowed on those masters who, by arduous discipline and training, come so close to the divine that they no longer need physical eyesight for their works of illumination, and so on.[114]

If the novel's readers feel any ethical value in Bihzad's act, it is due to Pamuk's skill as a storyteller. Cooke suggests that "a powerful – poetically gifted – writer/translator can enable

[113] Ibid, p 486-7.
[114] Ibid, p 486.

readers to recognise exemplary acts even in contexts that are not readily accessible to them," and that Pamuk succeeds in doing this. "Pamuk invites us into an interpretative context that is alien to us in many respects; yet he succeeds in making us feel sufficiently at home there to be able to experience the exemplarity of the figure of Bihzad, perhaps not with full force but at least to some degree." She concludes that ethical appropriation "is not simply a matter of rephrasing certain contents; rather, it has to be understood as a matter of creatively rearticulating the inspirational contents in a way that produces subjective experiences in which truth is manifested anew."[115] I will refer to this as *type two ethical appropriation.*

In the first type of ethical appropriation, the appropriators already share or understand the cultural context in which the appropriated content originates. When they appropriate the content, they import it into their ethical discourses whole cloth, along with its cultural context. In the second type, the cultural context of the appropriated content is foreign or remote to the appropriators. The ethical value which they wish to appropriate must be re-presented using techniques analogous to those of art and literature, so that it makes sense to them and retains its exemplary force. What is happening here, I suggest, is that the original context of the appropriated ethical value is being removed, and replaced with a new one which is familiar to the appropriators. This is in accord with Habermas' understanding of ethics. "Ethical knowledge," he says, "retains its capacity to provide orientation… only within the horizon of the established everyday practice of individuals socialized into a specific culture," in contrast to empirical scientific knowledge, which is contextless; in fact "we would destroy our ethical knowledge by submitting it to scientific examination, because theoretical objectification would dislodge it from its proper place in our life." (JA 22)[116] Ethical values must have a context; when modern subjects appropriate ethical values, they either come with their original contexts

---

[115] Ibid, p 487.

[116] Here Habermas is discussing Bernard Williams' view of ethical knowledge, via Aristotle, but the statement certainly fits his understanding of the ethics/morality distinction.

attached (type one appropriation) or are given ersatz contexts (type two appropriation). Whether modern subjects use type one or type two ethical appropriation to extract ethical values from an artwork or religious tradition will depend on how close they are to its cultural context.

Habermas' modest conception of postmetaphysical philosophy leads to a division of labour. Philosophy refrains from engaging in ethical discourse, while art and religion provide funds of "strong evaluations" which modern subjects can appropriate. Ethical appropriation is a consequence of the ethical abstinence of the postmetaphysical paradigm. It is not accurate to say, then, that Habermas' thinking was entirely secular before 2001. He pays little attention to religion, it is true, and assumes that the decline of religious and metaphysical worldviews is an uncontroversial feature of modernity. But by formulating an ethically abstinent conception of philosophy, he draws religion by default into the heart of modern ethical discourse. Ethical appropriation is the mechanism by which religion enters; drawing on Cooke's ideas, I have given an account of it in this section. I will make use of this account in my evaluation of Habermas' model of postsecular deliberative democracy in Part 3.

Of course, ethical appropriation is not the only way in which religious ideas come into contact with postmetaphysical thinking, as I will now explain.

2.6 Conceptual Appropriation

In his 1988 essay "Transcendence from Within, Transcendence in this World," Habermas says
that the history of philosophy is full of appropriations of religious concepts.

> Just as German Idealism with the concept of the Absolute appropriated
> theoretically the God of creation and gracious love, it also with a logical
> reconstruction of the process of the world as a whole appropriated theoretically
> the traces of salvation history. Also, Kant cannot be understood without recognizing
> the motive of conceiving the essentially practical contents of the Christian tradition
> in such a way that these could perdure before the forum of reason. (T 304)

Here we have three examples of appropriated concepts: the Absolute in German Idealism, which
is an appropriation of God; the world process in, presumably, Hegel and Marx, which is an
appropriation of salvation history; and some unspecified aspects of Kant's philosophy, which are
appropriations of "the practical contents of the Christian tradition."[117] The process Habermas is
describing here seems different to ethical appropriation. In all his examples it is carried out by
philosophers, for philosophers, with the goal of acquiring technical philosophical concepts.
Ethical appropriation, by contrast, can be carried out by any modern subject, and aims at the
acquisition of ethical values. I will call this kind of appropriation *conceptual appropriation.*

Concepts acquired in this way are evidently different from their religious originals – the
Absolute is not literally the Judaeo-Christian God. They are not simply the same concepts *tout
court*, restated with a different vocabulary. Something is carried across in a conceptual
appropriation, and something drops out. Habermas continues:

> The sublation of the world of religious representation in the philosophical concept
> enabled the saving of its essential contents only by casting off the substance of its
> piety. Certainly, the atheistic core, surrounded by esoteric insight, was reserved for
> the philosophers. (T 304)

An "atheistic core" is appropriated, and a "substance of piety" discarded. These terms need to
be clarified. Habermas' choice of words suggests that he conceives of conceptual appropriation

---

[117] In a later text, "The Boundary between Faith and Knowledge: On the Reception and Contemporary
Importance of Kant's Philosophy of Religion" (BNR 209-247), Habermas identifies the Sumum Bonum
and the idea of radical evil from Kant's late work as conceptual appropriations of elements of
Christianity. He thinks that the former is a more successful appropriation than the latter.

as a process which filters out the specifically religious aspects of a concept which originates in a religious tradition, and allows what remains to pass through into philosophical discussion.

It is *methodical atheism* which acts as a filter. Habermas asserts that "all philosophical appropriation of essentially religious contents" (T 304) are methodically atheist:

> methodical atheism is the manner of the philosophical reference to the contents of religious experience. Philosophy cannot appropriate what is talked about in religious discourse *as* religious experiences. These experiences could only be added to the fund of philosophy's resources, recognized as philosophy's own basis of experience, if philosophy identifies these experiences using a description that is no longer borrowed from the language of a specific religious tradition, but from the universe of argumentative discourse that is uncoupled from the event of revelation. (T 309)

In other words, when philosophy appropriates a religious concept, it *ceases to be* a religious concept in some way. This is the case even for metaphysical thinkers, and those who themselves have religious beliefs. (T 308-9) Habermas sees methodical atheism simply as part of the method of doing philosophy. It says nothing about the appropriating philosopher's own beliefs. Hegel, says Habermas, used methodical atheism, despite being a metaphysical thinker and a believer of sorts. This aspect of conceptual appropriation, in which methodical atheism filters out the "substance of piety" of a religious concept, bears a resemblance to type two ethical appropriation, in which the original cultural context of an appropriated ethical value is discarded.

Let us consider an example of conceptual appropriation. In his early essay "Between Philosophy and Science: Marxism as Critique," Habermas describes the appropriation by Hegel, Marx and others of an idea originating in the work of the 16th century Cabbalist Isaac Luria and the Protestant mystic Jakob Böhme. This is the idea of salvation history which is refashioned into the idea of the world-process, as referred to in the "Transcendence" essay. As Habermas puts it, the original idea is a

> remarkable conception of theogony and cosmogony, according to which the God of Origin… becomes external to Himself… by going into exile within Himself… [with the result that] man, left to himself alone in history to confront the work of his

redemption, must simultaneously carry out the redemption of nature with his own powers, and even the redemption of the overthrown God… [God] has surrendered Himself completely to the risks of an irretrievable catastrophe; it is only at this cost that He has initiated the world process as history. (TP 215)

So, the object of conceptual appropriation is the idea that God externalises himself in creation, goes into internal exile, so to speak, and by doing so turns history into a teleological process in which human beings must complete the task of restoring God to his eminence, and thus restoring harmony to creation. It functions as a theodicy, since it addresses the problem of evil. The reason a benevolent and omnipotent God allows evil to exist is that he has gone into exile within himself, within his own creation. God's internal exile makes history a teleological process, in which human beings bear the responsibility of bringing about a divine restoration.

Hegel appropriates this motif from Böhme, according to Habermas, refashioning it into his view of history as a dialectical process of the evolution of Spirit, from unreflexive unity through division into reflexive unity.[118] Habermas says that Hegel "plagiarizes" Luria and Böhme's "sophisticated myth" and then "rationalizes" it. (TP 215-6) Marx then takes up the idea from Hegel and uses it to structure his account of the relationship between living labour and dead labour. Living labour, dominated by the dead labour of capital, is like God exiled in creation; the proletariat's task of freeing itself from capital equates to humanity's task of restoring God to himself.

What can this example tell us about conceptual appropriation? What we need to determine is how the constraint of methodical atheism strips out the "substance of piety," and decide what, precisely, the remaining "atheistic core" of the appropriated concept is. When

---

[118] Böhme's ideas were well known to Hegel – see GWF Hegel, "Fragment on the "Life-Course" of God," in Jon Stewart ed., *Miscellaneous Writings of G. W. F. Hegel* (Evanston, Illinois: Northwestern University Press, 2002), pp 256-263, and the section devoted to him in GWF Hegel, *Lectures on the History of Philosophy* vol. III (New Jersey: Humanities Press, 1983), part 3 §1B (pp 188-216). For Böhme's influence on Hegel, see GA Magee, *Hegel and the Hermetic Tradition* (Ithaca and London: Cornell University Press, 2001), pp 36-50. Magee points out (p 38 n 61) that Habermas mentions Böhme in his dissertation on Schelling.

Habermas says that all philosophical appropriations of religious content are methodically atheist, he seems to mean that philosophy (metaphysical or postmetaphysical) cannot speak the language of any particular religion, a "specific religious tradition" coupled to "the event of revelation." (T 309) It can use concepts from a religion, as long as they are stripped of their connections to that context. He explains in the "Transcendence" essay that while theology must remain tied to the particular religion which it explicates, philosophy must remain open to (but at a distance from) all traditions, since its task is to mediate between (all) expert cultures and (all) sections of the lifeworld. Hegel and Marx can make use of Luria and Böhme's ideas, as long as the former's references to the particulars of Cabbalistic Judaism, and the latter's references to mystical Protestantism, are omitted. So, we can characterise the concept's substance of piety as its reference to the tenets of the religion from which it comes. This is what it must lose during a conceptual appropriation. The concept's atheistic core, then, is presumably that aspect of the concept which need not be connected to the tradition from which it originates. What might this be?

Consider Habermas' example again. All the various conceptual appropriations of Luria and Böhme's idea retain the structure of the original idea, but use this structure to perform different tasks or think about different phenomena. They operate in different registers, using different paraphernalia. The basic shape of the concept stays the same, but the components used to make up this shape change. Luria and Böhme talk about God and creation; Hegel talks about Spirit and the world-process; Marx talks about living and dead labour. The stories they tell have a single shape: there was an original unity, which externalizes itself and becomes divided, and is ultimately restored at a higher level. It is the bare structure of the idea which passes through the filter of methodical atheism, emptied of the contents of any particular religious tradition. I have used the term "motif" to refer to the object of appropriation. The same musical motif can be played on different instruments, in different registers, at different tempi. It is possible to say that two pieces of music display the same motif, even though every element

making up the motif is different. Similarly, we can say that Luria and Marx are using the same motif of thought, even though the former is thinking about God and creation and the latter is thinking about living and dead labour. Their very different theories share the structure of an idea, the shape of a concept. *This* is the atheistic core which philosophy appropriates from religion. If anything, thinking of this as a process of carrying something across from one side to another, as if the motif was a piece of cargo transported across a river, is something of a distortion. The atheistic core is not so much transferred as *replicated*, its shape rebuilt on the far shore.

So, conceptual appropriation can be defined as the extraction of motifs from religious traditions, through the filter of methodical atheism. These motifs can be rebuilt out of many different intellectual materials, and used to tackle different philosophical problems. Although Habermas draws no distinction between conceptual and ethical appropriation – I have coined the terms myself – the processes differ in a number of ways. Ethical values can only function within a context; they either bring their original context with them, or must be given a new one, when they are appropriated. Appropriated motifs, by definition, are stripped of their contexts. This is the only way they can become usable concepts for philosophy. Conceptual appropriation therefore transforms its appropriated content more radically than ethical appropriation. Ethical appropriation can be carried out by any modern subject; conceptual appropriation is a technique used by philosophers. Ethical appropriation occupies its position within Habermas' body of work as a result of his postmetaphysical paradigm of philosophy. Modern subjects, under conditions of postmetaphysical thinking, cannot look to Habermas' variety of philosophy for ethical values; as a result, he thinks that they must appropriate them from religion and art. This is how *Habermas* understands the ethical appropriation which is a feature of the modern age. He does not think of conceptual appropriation as being tied to the postmetaphysical paradigm; in fact, it seems to be widespread in the history of philosophy. Habermas mentions conceptual appropriations made by Kant, a subjectivist, and Hegel, a metaphysical thinker.

Other authors have identified examples of philosophers appropriating concepts from religion. Stephen Mulhall has written about the motif of the Fall and original sin in the work of Nietzsche, Heidegger and Wittgenstein, and Peter Dews about the concept of evil in the work of Kant, Fichte, Schelling, Hegel, Nietzsche, Schopenhauer, Levinas, and Adorno. Drawing on Jean Nabert's idea of "the unjustifiable," Dews defines evil as "a potential which is deeply rooted in the structure of human agency, and yet results in actions and processes that, from an ethical perspective, *absolutely should not be*."[119] He considers this motif, which originates in the Judaeo-Christian religious tradition, to have been replicated in the work of the philosophers he discusses. Similarly, Mulhall defines the concept of the Fall or original sin as "a conception of human nature as always already averting us from the relation to truth, comprehension and clarity that is nevertheless our birthright – hence, as structurally perverse or errant and yet redeemable from that fallen state –" but not by our own efforts, due to that very perversity.[120] Frankfurt School thinkers are well known for making use of religious concepts.[121] To cite the two most famous examples, Adorno uses the ban on graven images from the Second Commandment in his critique of identity thinking[122], and Benjamin uses the idea of the Messiah in his philosophy of history.[123] The consensus among commentators seems to be that appropriation changes a concept fundamentally. Dews says that "(i)nevitably, this drawing on religious sources involves a more or less radical reworking of theological concepts and motifs."[124] Mulhall says that the

---

[119] Peter Dews, *The Idea of Evil* (Oxford: Blackwell, 2008), p 9.

[120] Stephen Mulhall, *Philosophical Myths of the Fall* (Princeton, NJ: Princeton University Press, 2005), p 11.

[121] See Eduardo Mendieta ed., *The Frankfurt School on Religion*: *Key Writings by the Major Thinkers* (New York and Oxford: Routledge, 2005).

[122] Theodor W. Adorno, *Negative Dialectics* (New York and London: Continuum, 2007), p 207.

[123] Walter Benjamin, "Theses on the Philosophy of History," in *Illuminations* (See especially theses II, VI, XVII, XVIII, A and B.)

[124] Dews op. cit., p 119.

philosophers who conceptually appropriate the motif of original sin alter it in various ways. They all discard the notion that human beings can only be redeemed from outside the human world, and each one plays through the motif in a different way. "In Nietzsche, the central issue is the idea of our fallenness as punishment (more specifically, as self-punishment); in Heidegger, it is the idea of fallenness as embodiedness (more specifically, as mortality and animality); in Wittgenstein, it is the idea of fallenness as transgression (more specifically, as the refusal of limits and the perversion of desire.)"[125] In all these cases, a religious concept is stripped of its "substance of piety," reduced to the "atheistic core" of a motif, and redeployed by a philosopher.

I have investigated conceptual and ethical appropriation at length, partly to explore the curious position of Habermas' postmetaphysical paradigm. It eschews metaphysics, but makes use (like many other philosophies) of conceptual appropriation, and creates the conditions for ethical appropriation. It is as if the paradigm, by abstaining from the traditional concerns of ethics and metaphysics, compels its adherents to make use of these supposedly archaic ways of thinking. It turns away from metaphysics, but at the same time leans on it. I will return to this point in the Conclusion.

More concretely, the concepts of ethical and conceptual appropriation can be used as tools to evaluate Habermas' recent account of postsecular politics. His model of postsecular deliberative democracy, which I will analyse in Part 3, relies on a procedure of "translating" religious inputs to the public sphere into secular language. A great deal rests on this procedure, which Habermas never describes in detail. But with conceptual and ethical appropriation we have two modes, described in Habermas' writings, in which religious concepts can be taken over into non-religious discourses: philosophical discourse in the case of the former, modern ethical

---

[125] Mulhall op. cit., p 14.

discourse in the case of the latter. The procedures which I have excavated in the last two sections

will soon prove their worth.

2.7 The Anthropic Problem

So far I have described the peculiar relationship between religion and the paradigm of postmetaphysical thinking, as it was developed by Habermas during the 1980s and 1990s. I will now turn to the question of why he adopted his new postsecular stance in 2001.

Habermas says, and most commentators on his work agree, that this change of direction was driven by external factors: he gradually became aware that a secular political system places an unjustifiable cognitive burden on religious citizens (BNR 125-6); his engagement with Rawls in 1995 highlighted the problem of requiring religious citizens to "split their identities" into public and private segments; the terrorist attacks of September 11th 2001 demonstrated the danger of religious believers becoming alienated from liberal democracy.[126] (FHN 102-3) All of these factors are exogenous to Habermas' philosophical project. They are motives for change which originate outside his body of work. In the remainder of Part 2, I will argue that a major reason for the postsecular turn – in my view, the decisive reason – is endogenous. It was the inability of the postmetaphysical paradigm to deal with a particular problem which pushed Habermas into making his dramatic change of direction. I call this the *anthropic problem*. To understand it, we must look again at some of the basic assumptions of postmetaphysical thinking.

Recall Habermas' version of the ethics/morality distinction. Ethical discourse concerns questions of the good, and aims to clarify ethical values, which are specific to individuals and communities. Moral discourse concerns questions of justice, and aims to filter out valid norms, which apply to all moral subjects. The postmetaphysical element which Habermas adds to this distinction is the idea that philosophy should abstain from ethical discourse and contribute only

---

[126] As ever, Maeve Cooke is the exception to the rule. She lists Habermas' concern about genetic modification as one of the factors behind the postsecular turn, but does not connect it to the anthropic problem as I do. See Cooke, "Salvaging and Secularizing the Semantic Contents of Religion: The Limitations of Habermas's Postmetaphysical Proposal" *International Journal for Philosophy of Religion* 60 (2006), p 189.

to moral discourse. Ethical values are entangled with metaphysical claims about the good, human nature, and the telos of life or history. The principles which govern moral discourse – (D) and (U), in Habermas' theory – can be rationally reconstructed from the processes of communication immanent in the lifeworld, and are therefore quasi-transcendental and falsifiable. Just as it refrains from making metaphysical claims, philosophy which aims to be postmetaphysical should refrain from engaging in ethical discourse – this is Habermas' assumption.

But does postmetaphysical philosophy really avoid making use of metaphysical and ethical concepts? Does it, for example, abstain entirely from making metaphysical claims about human nature, and ethical ones about the good for human beings? Postmetaphysical philosophy's method, according to Habermas, is rational reconstruction. This method makes explicit the implicit knowledge contained in the species-wide capacities and cognitive structures of human beings. Habermas would say that his major theories of formal pragmatics, morality, and law and democracy are all based on rational reconstructions of the species-wide capacities of human beings. Now, to state the obvious, rational reconstruction can only reconstruct those capacities and stores of knowledge which (Habermas believes) the species actually has. He can extract formal pragmatics from the communicative processes of the lifeworld because human beings, as a matter of fact, are language users who make use of three kinds of validity claim connected to three worlds, attitudes, and grammatical persons. He can extract the discourse theory of morality from lifeworld procedures for repairing broken communication because human beings, as a matter of fact, are creatures who inhabit an intersubjective lifeworld and co-ordinate their actions using validity claims. These facts about human beings are the fodder for rational reconstruction, and all the theoretical products of Habermas' paradigm rely on members of the human species *being a certain way*. To this extent, postmetaphysical thinking does make use of an unstated conception of what it is to be human – a theory of human nature. At the very least we can say that Habermas, while doing philosophical work in his

postmetaphysical manner, assumes that a certain account of human nature is true. In "Remarks on Discourse Ethics," he refers to this as the human "form of life" or "life-form" (*Lebensform*). (JA 83)

On a strict interpretation, this is a violation of the tenets of postmetaphysical thinking. The paradigm should not endorse *any* theory of human nature, however plausible or persuasive, just as it should not endorse any account of the physical world, even that of the natural sciences, however well attested.[127] Habermas did not at first see the paradigm's reliance on an account of human nature as a problem. Consider this passage, in which he argues against Apel's strongly transcendental conception of the preconditions of discourse, and for his own quasi-transcendental a posteriori conception:

> This demonstration of the factual inescapability of substantive normative presuppositions of a practice *internally* interwoven with our sociocultural form of life is indeed conditional on the constancy of this life-form. We cannot exclude a priori the possibility of its undergoing changes. But that remains an empty possibility since – science fiction scenarios that transform human beings into zombies aside – we cannot imagine a fundamental alteration in our form of life. (JA 83-4)

Habermas is careful to say that the human form of life could in theory change. As a postmetaphysical thinker, he must not make any unfalsifiable claims. However, it is clear that he does not consider this a realistic possibility. His assumption seems to be, then, that his unstated theory of human nature is an accurate account of the invariant facts about human beings, and therefore a solid basis for making rational reconstructions. This may not be good postmetaphysical practice, but it does not involve the paradigm in any serious difficulties.

By the end of the 1990s, however, Habermas was not so sanguine about the immovability of the human form of life. Some of the "science fiction scenarios" which he had

---

[127] The difficulty which I am outlining here, then, has points in common with the difficulties which afflict Habermas when he, as a postmetaphysical philosopher, attempts to endorse any scientific theory. For example, he builds a certain interpretation of the theory of evolution into his account of "soft naturalism," which I discuss in the next section. Puntel claims that this draws him into making metaphysical assertions. See Puntel, "Habermas' Postmetaphysical Thinking," pp 33, 36-7.

once dismissed now seemed close to being realized. Two trends in particular seemed to Habermas to have the potential to alter the human life form: the spread of a "hard naturalist" worldview, and the development of technologies which enable the genetic modification of human beings, by human beings. Both these trends, according to Habermas, threaten to disrupt fundamental features of the human form of life, of the way our species understands itself. We have always seen ourselves as autonomous moral agents, responsible for the actions we perform. Under the influence of hard naturalism, however, we might begin to see our actions as the effects of chemical and electrical events taking place "behind our backs," so to speak, at the neurological level. The concepts of free will and moral responsibility would therefore be disrupted. We have always seen ourselves as naturally born beings, all members of the same lifeworld, able (in principle at least) to enter into symmetrical relations of communication with each other. Due to the development of biotechnologies, however, some human beings might see themselves as designed products, others as designers. The fundamentally egalitarian nature of the lifeworld, of intersubjectivity itself, might come to an end. This, at least, is what Habermas fears. I will go into greater detail about hard naturalism and genetic modification, and how Habermas thinks they threaten the human life-form, in the next two sections.

Such changes to human nature would, of course, wreak havoc with the unstated anthropic basis of Habermas' philosophical project. His postmetaphysical philosophy bases its theories of democracy, morality and communication on rational reconstructions of the *current* human form of life. It draws on the implicit knowledge and species-wide capacities of human beings as rational speaking and acting subjects, who inhabit an intersubjective lifeworld. If human beings are *not*, or are *no longer*, rational speaking and acting subjects who inhabit an intersubjective lifeworld, then formal pragmatics, the discourse theory of morality, and the discourse theory of law and democracy all collapse.[128] One would therefore expect Habermas to

---

[128] It is not just Habermas' work which is at stake. If he is correct about the corrosive effects of hard naturalism and genetic modification on the human form of life, then any moral theory which makes use of the concepts of moral responsibility or autonomy would be undermined.

argue against the changes to the human form of life which he expects hard naturalism and genetic modification to bring about. It is at this point, however, that postmetaphysical philosophy reaches a dead end, in the form of the anthropic problem.

Once it becomes clear that more than one human life-form is possible, then the question of which one is best for us arises. To safeguard the anthropic basis of his philosophical project, Habermas needs to argue that our current form of life is preferable to a form of life coloured by hard naturalism or genetic modification. But this is precisely what he cannot do, since *he considers the question of which is the best form of life for human beings to be an ethical question*, and as a postmetaphysical philosopher Habermas cannot participate in ethical discourses.

It is clear that Habermas considers this question to be ethical, rather than moral or pragmatic. His main text on the topic, which deals with the possible ramifications of genetic engineering, is called "The Debate on the Ethical Self-Understanding of the Species." (FHN 16-100) There, Habermas concedes that the debate opened up by the development of biotechnology has undermined his belief in the priority of moral norms over ethical values. He says that "the 'priority of the just over the good' must not blind us to the fact that the abstract morality of reason proper to subjects of human rights is itself sustained by a prior *ethical self-understanding of the species*, which is shared by all *moral persons*." (FHN 40) What I take Habermas to be saying here is this: he now understands that his discourse theory of morality is based on rational reconstructions of *a particular* form of human life. The issues raised by hard naturalism and genetic modification have made it clear that this form of life is contingent. We could adopt another. The continued existence of "the abstract morality of reason proper to subjects of human rights," as he has always understood it, depends on a *now-conscious decision* on the part of the human species to adhere to our original form of life. The ethical values

informing that decision are therefore prior to the norms produced by moral discourse: without them, that discourse could not take place at all. There is indeed an ethics underlying morality.[129]

Habermas seems to see the new discourse about how our species should understand itself, about which form of life is best for it, as a third variety of ethical discourse. Beyond the ethical-existential discourse of individuals and the ethical-political discourse of communities, we are now faced with what Maureen Junker-Kenny calls "species-ethical discourse." The issue of genetic modification, says Junker-Kenny, "forces [Habermas] to specify previously implicit assumptions on which the approach of discourse ethics depends and which are in danger of being eroded. This is why a moral argumentation is no longer deemed sufficient, and a new prior level, species ethics, is introduced."[130] Habermas classifies this kind of discourse as ethical, rather than moral, for a number of reasons.

Firstly, it is structurally similar to the discourses about the good for individuals and communities which I described in section 2.4. When we engage in species-ethical discourse we do not ask, "What is right in this situation? What principles could all affected parties agree to as just?" We ask, "What kind of species should we be? What is a good form of life, or the best form

---

[129] A number of commentators have argued that the priority of moral over ethical discourse in Habermas' thought is not sustainable. In *Insight and Solidarity,* which was written before 2001, Rehg considers this point, and concedes that the discourse theory of morality is sustained by a kind of ethics. The communicative practices of human beings can only be rationally reconstructed into a theory of morality if those human beings first believe in the value of open communication and cooperation. He points out that Habermas himself has never denied this. In "Morality and Ethical Life," Habermas says that "any universalistic morality is dependent upon a form of life that *meets it halfway*. There has to be a modicum of congruence between morality and the practices of socialization and education." (MCCA 207) Rehg argues, however, that "discourse ethics does presuppose the value of social cooperation based on mutual understanding, (but) this value is not just one conception of the good alongside other equally viable alternatives. Rather, the affirmation of this value draws on a solidarity so basic to being human that to deny it runs counter to rationality itself." See Rehg op. cit., p 8. But this saving throw no longer works, in the light of the debate opened up by hard naturalism and genetic modification. We can no longer be confident that the attributes which have hitherto been basic to being human are as stable as they appear. There are indeed viable alternatives.

[130] Maureen Junker-Kenny, *Habermas and Theology* (London and New York: T & T Clark, 2011), p 117. See also Maureen Junker-Kenny, "Genetic Enhancement as Care or as Domination? The Ethics of Asymmetrical Relationships in the Upbringing of Children" *Journal of Philosophy of Education* 39:1 (2005).

of life, for human beings?" These questions would draw on "strong evaluations" of what it is to be human. This is an ethical discourse on the grandest possible scale, admittedly, but it is an ethical discourse nonetheless.

Secondly, to regard the issue of which is the best form of life for our species from the moral point of view would beg the question. The issue at stake is whether the species should retain the form of life or self-understanding which has hitherto grounded its morality. An attempt to bring moral norms to bear on this issue would be circular. The species-ethical question, says Habermas, is "why... should we *want* to be moral? An assessment of morality as a whole is not a moral judgement, but an ethical one, a judgement which is part of the ethics of the species." (FHN 73)[131] Writing specifically about genetic modification, in the "Postscript" to "The Debate on the Ethical Self-Understanding of the Species," Habermas notes that

> "if eugenic manipulation changes the rules of the [moral] language game itself, this act can no longer be criticised according to those rules. Therefore, liberal eugenics provokes the question of how to value morality as a whole... No arguments from the moral language game itself can be mustered against a eugenic self-instrumentalization of the human species which changes the very rules of the game." (FHN 92)

We cannot adduce moral reasons for being the kind of species that has morality. The only way of dealing with the question of the best form of life for human beings is through ethical discourse.

Thirdly, the nature of the debate about the self-understanding of the species prevents participants from engaging in moral discourse. To examine a question from the moral point of view, according to Habermas, we must abstract away from particular contexts and life-histories, and cast the question in the formal terms of the (D) and (U) principles.[132] This is not always

---

[131] There is a parallel between this metaethical problem and that of the justification for bringing about the conditions of justification, which Habermas discusses in "Morality and Ethical Life" – see MCCA 208-9.

[132] It should be noted that Habermas also discusses *application discourses*, which deal with how to apply the norms arrived at in moral *justification discourses* in concrete situations. He borrows the idea from Klaus Günther. In application discourses, the *principle of appropriateness* replaces the (U) principle. (JA 37-8, 128-30) The fact remains that discourses of moral justification must be abstract and contextless.

possible. Habermas thinks that some practical questions are framed in terms which overlap with the normative self-descriptions and situation-descriptions of the persons asking the questions. This makes it impossible for them to abstract away from their context. If it is not possible to adopt the moral point of view, then these questions can only be addressed with ethical, not moral, discourse. Habermas considers the debate about abortion to be an example:

> It might transpire that descriptions of the problem of abortion are always inextricably interwoven with individual self-descriptions of persons and groups, and thus with their identities and life projects. When an internal connection of this sort exists, the question must be formulated differently, specifically in ethical terms. Then it would be answered differently depending on context, tradition and the ideals of life. (JA 59-60)

In the case of abortion, the situation-descriptions involved in framing the question – accounts of when a human being begins and ends, where one person's body becomes another's – are the same as the self-descriptions of the persons asking the questions – this is what makes me a human being, this is what makes my body my own. The participants cannot simply put aside their ethical self-understandings and adopt the moral point of view.[133] The obstacles to moving to the moral level explain why the issue of abortion remains so deeply controversial. It is impossible to arrive at a single, universally valid moral norm; all we have are a myriad of contradictory ethical values. Questions about the best form of life for human beings, or how we should understand ourselves as a species, evidently overlap with the self-descriptions and situation-descriptions of the human beings asking the questions; these questions must therefore be answered at the ethical rather than the moral level.

For all these reasons, Habermas thinks that the discourse about the best self-understanding or form of life for our species, which has been opened up by hard naturalism and

---

[133] Aside from this passage, Habermas says little about this particular obstacle to moving from ethical to moral discourse. The threshold for this problem – the point at which a practical question is so entangled with the self-descriptions and self-understandings of those asking it that they cannot adopt the moral point of view – must be quite high, however. *All* practical questions could be said to contain situation-descriptions which overlap with self-descriptions to some extent. If the threshold was set too low, it would rule out moral discourse as such.

genetic modification, must be an ethical discourse. "Where we lack compelling *moral* reasons,"

he says, "we have to let ourselves be guided by the signposts set up by the *ethics* of the species."

(FHN 71) Postmetaphysical philosophy, of course, can contribute only to moral, not ethical,

discourses. The contours of the anthropic problem have now become clear.

Let me recap the arguments which I have made in this section. To begin with, I argued

Habermas' postmetaphysical paradigm is dependent on an unstated account of human nature,

and that by the end of the 1990s Habermas was aware of this. The emergence of hard naturalism

and biotechnology suggested to Habermas that human nature, the human form of life, could

change. Such changes would undermine his entire philosophical project (along with much else),

as I will explain in the following sections. The only way to counteract these changes would be to

engage in a debate about which form of life is best for the human species. But, Habermas

believes, any debate along these lines can only take place at the level of ethical discourse. Finally,

Habermas continues to maintain that as a postmetaphysical philosopher, he must abstain from

all ethical discourses. (FHN 1-3)

This is what I mean when I talk about the anthropic problem. The problem is not that

postmetaphysical thinking tacitly relies on a theory of human nature; the problem is not that

human nature is less stable than we have previously assumed. The problem is that

postmetaphysical thinking, while relying on a contingent ethical self-understanding of the

species, disbars itself from participating in species-ethical discourse. The ethically-abstinent

paradigm cannot defend itself against threats to its own anthropic basis. In the face of potential

changes to the ethical self-understanding of the species, it is helpless. Habermas is caught in a

trap of his own devising. He can see the corrosive effects which hard naturalism and genetic

modification might have on our current life-form, and by extension on his philosophical project;

but the tenets of that philosophical project prevent him from engaging in the species-ethical

discourse which might save it.

I will argue in section 2.10 that the anthropic problem was the crucial endogenous factor behind Habermas' postsecular turn in 2001. Before reaching that point, however, I must explain several points which I have dealt with briefly in this section. I have not yet explained in any detail *how* hard naturalism and genetic modification, according to Habermas, might alter the ethical self-understanding of the species. Nor have I justified my claim that the postmetaphysical paradigm, ethically constrained as it is, is unable to tackle these problems. In the next two sections, then, I will explain the threats posed by hard naturalism and genetic modification, and evaluate Habermas' attempts to counter them using postmetaphysical resources.

2.8 Hard Naturalism

What does Habermas mean by *strong* or *hard naturalism*, and why does he consider it a threat to the ethical self-understanding of the species?

Naturalism, simply stated, is the belief that everything which exists is part of a single natural world.[134] Habermas calls this basic naturalist assumption *ontological monism*. (LG 14) It implies that everything which exists is subject to the same universal laws: an ontological monist would not accept that there are, for example, separate realms of matter and spirit which operate in different ways.[135] Habermas accepts ontological monism, and assumes that all modern thinkers would do the same.[136] What he objects to is the leap from a reasonable ontological monism to an unwarranted *epistemic monism*. Hard naturalists, according to Habermas, assume that since there is a single natural world, there must be a single valid way to describe and explain everything within it. From this they conclude that the scientific description of nature is the only correct one, and that all phenomena – including elements of the human world such as morality, rationality and free will – must be understood in natural-scientific terms, meaning causal, nomological terms. Habermas associates hard naturalism with thinkers such as Quine, (TJ 23) and with eliminative materialist or determinist approaches in the philosophy of mind.[137] He sees

---

[134] Yates op. cit, p 47.

[135] Both materialists and idealists could accept ontological monism, although they would disagree about what this unified nature consisted of – matter or mind. Hegel, for example, is an idealist ontological monist. Thinkers such as Kant and Descartes, on the other hand, are ontological dualists – they posit separate but interacting realms of noumenon and phenomenon, or mind and matter.

[136] By rejecting ontological dualism and implicitly endorsing ontological monism, is Habermas making a metaphysical assertion and violating the strictures of postmetaphysical thinking? He does not say explicitly that ontological monism is true, only that "scientifically enlightened persons" must assume that it is the case. (LG 14) He does explicitly reject ontological dualisms such as that of Kant. (LG 37) Dieter Henrich, for one, thinks that Habermas is trying to have his metaphysical cake and eat it, combining contradictory elements of naturalism and anti-naturalism. See Henrich "What is Metaphysics – What is Modernity?" pp 302-3.

[137] Wolf Singer is a frequent target of Habermas' criticism. See "Freedom and Determinism," (BNR 151-180), passim. Philosophers of mind such as Patricia Churchland and sociobiologists such as E O Wilson could be added to the list.

Wilfrid Sellars' famous thought experiment about the objectification of human behaviour as a *reductio ad absurdum* of hard naturalism, rather than an argument for it. (FHN 106-7)[138]

As I have noted, Habermas thinks that hard naturalism presents a threat to the current ethical self-understanding of the species. It endangers our understanding of ourselves as beings who have free will, act for reasons, and can be held accountable for our actions. Hard naturalism "allies itself with a scientific understanding of our cognitive abilities. All cognition is ultimately reducible to empirical processes." (TJ 23) Similarly, it reduces the intersubjective world to the objective: "A 'strongly' naturalistic explanatory strategy aims to *replace* the conceptual analysis of practices of the lifeworld with a scientific neurological or biogenetic explanation of the achievements of the human brain." (TJ 27) The upshot of this is that hard naturalism re-describes phenomena at the human level, which have hitherto been understood in terms of reasons and norms, in terms of causal chains and physical laws. Such reductive re-descriptions, according to Habermas, could undermine the "language game of responsible agency." To a hard naturalist, human beings do not act freely, for reasons, such that they can be held morally responsible for their actions. This is mere "folk psychology." Instead, their actions are determined by physical causes, ultimately describable in terms of electrical and chemical events in the brain. A human being is no more responsible for performing an action than a tree is responsible for bending in the wind. "As soon as [rational action] is replaced with the causality of nomologically determined events," says Habermas, "the phenomenon [of free will] disintegrates—even just on grammatical grounds… Without its two pillars—the presupposition of the causal effectiveness of the mind and the connection between reflection and freedom — the language game of responsible agency collapses." (LG 19-20) Hard naturalism thus downgrades the giving of reasons and justifications, the source of valid norms for Habermas, to a mere epiphenomenon of physical states. (BNR 163-4)

---

[138] See Wilfrid Sellars, "Philosophy and the Scientific Image of Man," in *Science, Perception, and Reality* (Atascadero, CA: Ridgeview, 1991).

In formal pragmatic terms, what hard naturalism does is to reduce the second person to the third person perspective. As I explained in Part 1, Habermas associates the grammatical second person with the intersubjective world, the norm-conformative attitude, and the validity claim of moral rightness. It is by mastering the second person perspective that human beings learn to regulate their shared social world through morality and law. He associates the third person perspective with the objective world, the objectivating attitude, and the validity claim of truth. Science and technology operate in this instrumental-rational area. (TCAI 238-9) This, ultimately, is the nature of the threat posed by hard naturalism to the current human form of life. If the hard naturalist worldview were to be generally accepted, the second-person perspective would be displaced by the third-person – the intersubjective world would collapse into the objective. All the distinctively intersubjective phenomena of the human world would either come to an end, or be transformed beyond recognition. The collapse of the moral language-game is only the most obvious example. Needless to say, Habermas' postmetaphysical philosophical project would be thrown into disarray by such a change. The phenomena which it rationally reconstructs, the bases of its theories, would no longer exist.

Habermas considers the hard naturalist worldview to be gaining influence both in academia and society as a whole, fuelled by recent scientific discoveries. (BNR 151, LG 14) He suggests that it has already had a concrete impact on the legal system, where causal explanations for behaviour, such as genetic predispositions, have started to displace the notions of criminal culpability and guilt. (LG 22-3) He fears that a legal system based on hard naturalist principles would simply regulate criminal behaviour with rewards and punishments, rather than holding individuals accountable for their actions. In other words, it would regress from postconventional to preconventional morality. In this way, Habermas thinks, hard naturalism has already started to alter the ethical self-understanding of the species.

Evidently, Habermas needs to argue against hard naturalism, and in favour of his alternative theory of *soft* or *weak naturalism*, which I will discuss briefly below. It is at this point, however, that the anthropic problem makes its presence felt. Habermas cannot simply assert that the one theory is false and the other true, or that one ethical self-understanding of the species is better for us than another. If he is to remain within the limits of postmetaphysical thinking, he must walk a narrow path between making metaphysical or ethical claims on the one hand, and ceding territory to hard naturalism on the other. His arguments against hard naturalism must not cross the line into metaphysics. Let us evaluate them.

Firstly, Habermas argues that hard naturalism is covertly metaphysical. By insisting that the causal, nomological descriptions and explanations which it associates with natural science are the only correct way to understand reality, it inflates the observer's perspective into an extramundane "view from nowhere." (LG 27) In this sense hard naturalism is closer to metaphysics than to modern science, or postmetaphysical philosophy.[139] "The ontologization of natural scientific knowledge into a naturalistic worldview reduced to 'hard' facts is not science but bad metaphysics." (BNR 207) Habermas thinks that the tension between ontological monism and our self-understanding as human beings can tempt us to retreat into metaphysics in two ways. The "idealistic fallacy" is to extend the epistemic dualism between the second and third person perspectives into an ontological dualism – to rescue free will and morality by returning to Kant's noumenal and phenomenal worlds. "Conversely, the naturalistic fallacy is but the other side of the same coin; it simply assimilates transcendental conditions to empirical conditions, without considering the aporia of self-referentiality, and projects them onto a scientifically objectified realm." (TJ 28) Hard naturalism reverts to metaphysics by inferring epistemic monism

---

[139] Related to this argument is Habermas' claim that hard naturalism starts from a caricatured understanding of natural science – it is not a scientific theory, but a "worldview based on a speculative interpretation of scientific knowledge." (BNR 152) For example, hard naturalists overestimate the extent to which scientific theories fit all phenomena into causally closed systems. Habermas points out that scientific findings are, in reality, far "messier" than this. (LG 38)

from ontological monism. To be properly postmetaphysical thinkers, says Habermas, we must reject it.

This argument is reasonable as far as it goes, but it would only be persuasive to those already committed to postmetaphysical thinking. Hard naturalists might not share Habermas' assessment of metaphysics. They could concede that hard naturalism is a form of metaphysics, and argue that it is the one *correct* metaphysical theory, which accurately describes the universe. This argument is safely postmetaphysical, but it is unlikely to counteract hard naturalism.

Secondly, Habermas questions the plausibility of what hard naturalists aim to achieve. Their mistake is to assume that the third person perspective is all-important, and that it is capable of replacing the second person perspective. "If naturalistic research programs of this kind were successful," says Habermas, "they would make it possible to replace the phenomena that are accessible from the participant perspective with objectifying self-descriptions." (BNR 203) But in fact, it is *not* possible for human beings to understand ourselves in third person terms. We cannot do without rationality, normativity, and "folk psychology." As long as we are speaking and acting subjects, we cannot recognise ourselves in the objectifying descriptions of hard naturalism. (TJ 24) "The limits of naturalistic self-objectification are trespassed when persons describe themselves in such a way that they cannot recognize themselves as persons any more." (LG 25) Habermas' argument seems to contain two claims. First, he asserts that phenomena which reside within the participant's perspective cannot be redescribed from the observer's perspective without loss of meaning; and second, that human beings cannot function if they understand everything, including themselves, from the observer's perspective.

Habermas makes use of other philosophers' arguments to defend the first claim. Thus, he suggests that Husserl and Frege showed that the intentional attitudes of persons cannot be redescribed as spatiotemporal events linked by cause and effect, while Wittgenstein showed that semantic content resides in symbolic expressions – no description of its physical substrate,

however detailed, can give access to it. (LG 25) "Thoughts that can be expressed in mentalistic terms cannot be translated without loss of meaning into an empirical vocabulary geared towards things and events." (BNR 167) Habermas' argument is that the kind of redescription which hard naturalists are attempting would sacrifice all that is significant and meaningful about the phenomena which they are attempting to redescribe. This appears to be safely postmetaphysical. Once again, however, it is not clear how effective the argument would be in countering hard naturalism. It would certainly be persuasive to someone who considers our traditional accounts of intersubjective phenomena to be valuable. But a hard naturalist could argue that these traditional accounts are simply false – they are pieces of naïve folk psychology, and should be discarded for the sake of attaining an accurate picture of reality. Any meaning lost through such a redescription is well lost.

Habermas' second claim, that human beings cannot function if they understand themselves in objectivating terms, can be made in a strong or a weak version. The strong version states that we *cannot* reduce our second person perspective to the third person. The weak version simply asserts that as socialized human beings, we either could not *imagine*, or would not w*ant* to do away with the second person perspective and all that goes with it. (LG 37)[140] There are problems with both versions of Habermas' claim, as Richard J. Bernstein points out. "There is strong interpretation that it is a 'conceptual impossibility' to reduce or translate the language of rational explanations into the language of causal explanations."[141] But this appears to be a transcendental, a priori, and thus metaphysical claim. As a postmetaphysical thinker, Habermas should restrict himself to falsifiable, quasi-transcendental assertions.

---

[140] Habermas is talking about Sellars' thought experiment in this passage, but the same can be said about his own position, as Bernstein's argument shows.

[141] Richard J. Bernstein "Naturalism, Secularism and Religion: Habermas' Via Media," *Constellations* 17:1 (2010), p 159.

The weaker interpretation is that "given our current knowledge, we cannot make sense of eliminating either of these epistemic perspectives."[142] This claim is compatible with the postmetaphysical paradigm. The problem with it is that as an argument against hard naturalism, it misses its mark. It may not be possible for human beings *as we currently exist* – as beings with a certain form of life and ethical self-understanding – to imagine reducing the second person perspective to the third. But the entire point of the debate opened up by hard naturalism and genetic modification is that we may not always be the kind of beings we are today. Under the influence of hard naturalism, we may become entirely different creatures, who *could* operate in third-person social world. As Bernstein puts it, we "cannot conceptually rule out the *possibility* of the 'success' of a sophisticated version of hard naturalism. At best, we can give good reasons to show why this is highly unlikely."[143] The upshot of the weak version of Habermas' argument is simply that the ethical self-understanding of the species which would follow from hard naturalism is radically different to the one we currently have. This is true enough, but it is not an argument for choosing one ethical self-understanding over the other.

Thirdly, Habermas argues that the reductivist programme of hard naturalism cannot succeed because it involves a performative contradiction. Hard naturalism takes the scientific description of the world as its guiding light. Intersubjective lifeworld practices, hitherto understood in participatory or second person terms, must be redescribed in third person terms, on the assumption that the only valid kind of explanation is causal and nomological. The problem with this, as Habermas points out, is that science is *itself* an intersubjective lifeworld practice. (LG 38) The production of scientific knowledge involves more than observation – it relies on the community of scientists offering and criticising reasons, engaging in discourse about data and theories, and so on. These activities can only take place intersubjectively, from the participant's

---

[142] Ibid.

[143] Ibid.

perspective. (BNR 169-70) Although its outputs may operate in third person terms, the practice of science itself rests on scientists regarding each other as participants in discourse, from the second person perspective. For this reason, argues Habermas, hard naturalism "is already thwarted by the fact that investigation of the objective world itself rests on an argumentative dispute, which, although it appeals to events accessible from the observer perspective, is nourished by hermeneutic resources." (BNR 207) Hard naturalism cannot give an account of its own standards of valid explanation and description in hard naturalist terms – it is caught in a performative contradiction.[144]

This, it seems to me, is Habermas' most successful argument against hard naturalism. One the one hand, it does not violate any of the strictures of postmetaphysical thinking; on the other, it seriously undermines the claims of hard naturalism. So despite the difficulties which I have reviewed above, Habermas has at least one argument against hard naturalism which is both powerful and safely postmetaphysical.

In opposition to hard naturalism, Habermas proposes his alternate theory of soft naturalism. Its central tenet is a combination of ontological monism with the *epistemic dualism* of the second and third person perspectives. (BNR 206) The theory accepts the scientific account of the world, while at the same time aiming to give intersubjective phenomena such as rationality and morality their due. Habermas believes that he can argue for soft naturalism without crossing the line into metaphysics, a claim which many of his critics would dispute.[145]

---

[144] Habermas has been making versions of this argument for a long time, although its original target were logical positivism and "scientism" rather than hard naturalism. See his essay "The Analytical Theory of Science and Dialectics" in *The Positivist Dispute in German Sociology,* especially pp 151-4.

[145] Aside from the criticisms by Puntel and Henrich which I have already mentioned, consider these comments from Randolph Clarke: "Despite his invocation of postmetaphysical thinking, Habermas seems ready to embrace some heavy-duty metaphysics when he recommends what he calls weak naturalism. The proffered picture is one of a multi-layered world containing both physical and weakly (and perhaps also strongly) emergent properties, with nomic causation as well as an anomic causality of reasons, and causal lines running upward as well as downward. Though reasons are said to reside within nature, they are consigned to a sub-realm impenetrable by natural science. Whatever the warrant for these extravagant conjectures (and I am skeptical of its strength), the view seems unlikely to accomplish

Soft naturalism has its own strengths and weaknesses, but it is not necessary for me to discuss it further. What I have done in this section is to show how Habermas considers hard naturalism a threat to the current ethical self-understanding of the species, and how he attempts to argue against it without violating the tenets of postmetaphysical thinking. This is a difficult task. Habermas cannot simply argue that our current species-ethic, in which we understand ourselves to have free will and moral accountability, and to act for reasons, is better for us than the species-ethic of hard naturalism. This would be an ethical argument. Nor can he argue that hard naturalism gives a false picture of reality. This would be a metaphysical argument. His strategy for defending the ethical self-understanding of the species from hard naturalism must be far more indirect. It seems to me that Habermas makes one argument to this effect which is both postmetaphysical and effective: that hard naturalism involves a performative contradiction. He just about succeeds, then, in countering the hard naturalist threat to the ethical self-understanding of the species. It remains to be seen whether he will be so successful in dealing with genetic modification.

---

any more than can a much more austere one when it comes to saving the appearance of freedom." Randolph Clarke, "The Appearance of Freedom," *Philosophical Explorations* 10:1 (2007), p 55.

To understand the postsecular turn, we must first understand the circumstances under which the seriousness of the anthropic problem became obvious to Habermas. Hard naturalism presents a threat to the ethical self-understanding of the species, but, as we saw in the previous section, Habermas finds a way of arguing against it without departing from the postmetaphysical paradigm: he argues that hard naturalists are caught in a performative contradiction. The possibility of the genetic modification of human beings, by human beings, poses a similar threat; but in this case, I will suggest, Habermas was unable to find any postmetaphysical means of arguing against it. Unable to intervene in the species-ethical discourse opened up by genetic modification, Habermas was forced to reach outside the postmetaphysical paradigm, for ethical and metaphysical concepts which could be used to defend the ethical self-understanding of the species from the threat of biotechnology. Ultimately, he reached for religion. My argument is that the issue of genetic modification, combined with the anthropic problem, was the crucial endogenous factor behind Habermas' postsecular turn in 2001.

But what evidence is there that Habermas was concerned about genetic modification, and the inability of postmetaphysical thinking to address it, in the period leading up to 2001? To answer this question, we must look back to a bitter public controversy which took place in Germany in the summer of 1999.

At Schloss Elmau in Bavaria in July 1999, Peter Sloterdijk presented a paper to a conference on Heidegger. "Rules for the Human Zoo"[146] (*Regeln für den Menschenpark*) begins as a commentary on Heidegger's "Letter on Humanism." In it, Sloterdijk gives a literary or epistolary definition of humanism. Books are "thick letters to friends,"[147] open-ended messages

---

[146] Peter Sloterdijk, "Rules for the Human Zoo: a response to the *Letter on Humanism*," trans. Mary Varney Rorty, *Environment and Planning D: Society and Space* 12 (2009) pp 12-28.

[147] Sloterdijk, p 12.

sent to strangers in the hope of establishing friendly communication. They are benevolent chain-letters between generations and cultures. The culture of Western humanism first arose via the letters the Greeks sent to the Romans, which were sent on in turn to early modern Europe. This process is not an end in itself. Reading great literary and philosophical texts establishes friendship between distant people, and creates communities. Traditionally, according to Sloterdijk, you become a full member of the nation through military service and reading the national canon; you become a cosmopolitan humanist through reading the international canon.[148] This reveals the ultimate goal of humanism – to tame and domesticate human beings. It is through *Bildung*, the process of education and acculturation, that we civilize ourselves. There has always been a struggle, says Sloterdijk, between the civilizing and bestializing tendencies within human beings. Humanism, in the form of reading thick letters from friends, and thus creating bonds of friendship and community, was a method for reigning in our bestial nature. Sloterdijk offers the Roman Empire as the exemplar of both tendencies – contrast the classical humanist in his study, surrounded by books, with the crowds in the arena, baying for blood. Each tendency has its own medium – books on the one hand, mass spectacle on the other.[149] "The latent message of humanism, then, is the taming of men. And its thesis is: reading the right books calms the inner beast."[150]

But this kind of humanism, says Sloterdijk, is no longer possible. The medium of humanism, the book, is obsolete. New media have arisen, from radio to the internet, and these new media are "postliterary, postepistolary, and thus posthumanistic." Books can no longer perform their civilizing, community-forming function, and so "it is no longer possible to retain the illusion that political and economic structures could be organized on the amiable model of

---

[148] Sloterdijk, p 14.
[149] Sloterdijk, pp 15-16.

[150] Sloterdijk, p 15.

literary societies."[151] However, the problem which humanism tried to solve – our equal propensity for barbarism and civilization – remains. "What can tame man," asks Sloterdijk, "when the role of humanism as the school for humanity has collapsed?"[152] If our books cannot domesticate us anymore, what can?

Sloterdijk turns to Heidegger for a possible alternative. Heidegger seems at first to occupy a very different position. He vociferously rejects the idea that humans are animals plus reason, or plus culture – an idea which is integral to epistolary humanism. Humans and animals are in fact fundamentally, ontologically different, and no amount of reading can turn one into the other. As Heidegger argues in his *Fundamental Concepts of Metaphysics,* humans can conceive of their environment as a *world*, possessing objective existence. Animals cannot. They are "poor-in-world," trapped in their environment.[153] What is distinctive about humans is that they can become the "shepherds of Being," which they encounter in language, the "house of Being." This seems to avoid the whole problem of humanism – there is no need to worry about whether humans will become tame or bestial, if there is an ontological gulf between men and beasts.

Sloterdijk argues, however, that Heidegger is simply replicating epistolary humanism at a more fundamental level. The basic premise of this kind of humanism is that reading great books, communicating with distant members of the community, curbs our barbarism. Now, consider Heidegger's shepherds of Being. They must be quiet, receptive, and deeply involved in language (since this is where Being dwells). This is just as much of a taming process as any humanistic *Bildung*. As Sloterdijk puts it, "by describing man as the shepherd and neighbour of

---

[151] Sloterdijk, p 14.

[152] Sloterdijk, p 20.

[153] Martin Heidegger, *The Fundamental Concepts of Metaphysics: World, Finitude, and Solitude* (Bloomington: Indiana University Press, 1996)

Being and calling language the House of Being, [Heidegger] bound man into a relationship with Being that imposed radical constraints on his behaviour... man must become more passive, and tamer, than the humanist reading the classics."[154] Again, language and communication are central to this self-taming, albeit in the form of allowing Being to speak through our language, rather than exchanging thick letters with distant friends.

So Heidegger is, in many ways, attempting to accomplish the same task as the classical humanists, and with similar tools. There is no solution to the problems of humanism here. At this point Sloterdijk starts to bring in ideas from his own work. He argues that there is a long pre-history to human beings, which Heidegger (and others) ignore. We underwent an epochal process of transformation *before* we entered Heidegger's "Clearing," and became capable of listening to Being. There are two strands to this process, which I will call *natality* and *domesticity*.

Natality refers to the fact that humans are not born fully-formed. We start off immature and animalistic, and have to be inducted into humanity by learning language. It is at this point that we escape from our environment, and become capable of having a world. Sloterdijk calls this a second birth. (Here Sloterdijk is drawing on ideas from his magnum opus, the three-volume *Spheres* (*Sphären*).[155] He has always considered the mammalian nature of humans, the fact that we experience the womb followed by live birth, to be of fundamental importance.)

Domesticity refers to the fact that humans do not just enter into language as the House of Being – we also build literal, physical houses to inhabit, as well as other structures (Sloterdijk mentions temples and palaces). Once there is a house, there is an inside and an outside, and

---

[154] Sloterdijk op. cit., pp 18-19.

[155] Peter Sloterdijk, *Sphären I – Blasen* (Suhrkamp, 1998), *Sphären II – Globen* (Suhrkamp, 1999), *Sphären III – Schäume* (Suhrkamp, 2004).

decisions have to made about who belongs where. Once the "biopolitical unit" of "man, the house and animals" emerges, masters and pets exist for the first time.[156]

This links "Rules for the Human Zoo" to Sloterdijk's longstanding project of drawing attention to the spatial and physical nature of human beings, and the bodily basis of our thinking. He described the *Spheres* as a companion volume to Heidegger's *Being and Time, Being and Space.* With regard to the question of humanism, he is arguing that there are in fact *two* processes of humanization or taming: *Bildung* and *Züchtung*, reading and breeding. Epistolary humanism, and even Heidegger's ontological humanism, are focused exclusively on *Bildung*, culture, reading. This has now come to an end, due to the decline of its medium, the book. What remains for us, if we still want to avoid bestialisation, is *Züchtung*, cultivation, breeding. The shaping of human nature not via language, but via physical intervention. "(A)t no time was it, nor will it be, possible to accomplish the taming and befriending of men with letters alone... breeding, whatever form it may have taken, was always present as the power behind the mirror."[157]

Having made his point about taming humans via *Züchtung*, not *Bildung*, Sloterdijk turns to Nietzsche's *Zarathustra* and Plato's *Statesman* for guidance on how such breeding programmes should be run. What he takes from Nietzsche is the idea that more than one breeding programme is possible, and that we should face the choice between programmes openly. Zarathustra expresses contempt for the Last Men – "Virtue for them is whatever makes one modest and tame: with that they have made the wolf into a dog and the human being itself into the human's best domestic animal."[158] The Last Men are the result of a particular form of a taming, a programme of *Züchtung* aiming at harmlessness. Nietzsche regards the Judaeo-

---

[156] Sloterdijk, p 21.
[157] Sloterdijk, p 23.

[158] Friedrich Nietzsche, *Thus Spoke Zarathustra*, trans. Graham Parkes (Oxford: Oxford University Press, 2005), p 147.

Christian breeding programme, which hid behind the veil of epistolary humanism, as an attempt to inculcate slave morality in humans, all the more pernicious because it was done covertly. In fact we have a choice, between a minimizing breeding programme, and a maximizing one.[159] It is irresponsible to shirk this decision. We must be honest, and choose the direction of our species with our eyes open.

It is in this context that Sloterdijk makes his few references to genetic modification. We will soon have the "anthropotechnology," as he puts it, to intervene in our own physical make-up.[160] As such we will be compelled by circumstances to take on the role of self-selectors and self-breeders. We need to formulate a "codex of anthropotechnology"[161] to guide our decisions – saying "laissez faire" is not good enough. This codex is what Sloterdijk means by the "rules for the human zoo." His term "*Menschenpark*," which could be translated as "human zoo" or "human park" – Andrew Fisher in *Radical Philosophy* renders it as "human theme-park"[162] – has shock-value, but it is an idea drawn from the *Spheres*, and not an automatically authoritarian one. Sloterdijk believes that as immaturely-born mammals, we re-create the "micro-sphere" of the womb around us, in social, political and physical structures, which he refers to as "macro-spheres."[163] We voluntarily build macro-spheres around ourselves at the level of the household, the city, the nation-state, the continent, and so on. Sloterdijk says that "(h)umans are self-

---

[159] Sloterdijk does criticize some of the specifics of Nietzsche's argument – the timescale is far too short, and it is unlikely that there was ever a conscious conspiracy among priests and teachers. "(A)gentless biocultural drift" (Sloterdijk, p 23) is more probable. And he is careful to say that Nietzsche's Nazi admirers completely failed to understand the nature of the overman, and what Nietzsche meant by breeding for maximization.

[160] It is notable that the issues which attracted so much attention in the German public sphere in the following months – genetic engineering, cloning, and eugenics – are barely mentioned in the paper itself.

[161] Sloterdijk, pp 23-24.

[162] Andrew Fisher, "Flirting with Fascism – the Sloterdijk Debate," *Radical Philosophy 99* (Jan/Feb 2000), pp 58-60.

[163] See Sloterdijk op. cit.

fencing, self-shepherding creatures. Wherever they live, they create parks around themselves."[164] The only difference with Sloterdijk's post-humanist human zoo is that the issue of *Züchtung* is out in the open. The rule-book for the human zoo, the codex of anthropotechnology, regulates how we are to breed and domesticate ourselves.

This raises the issue of political power, of who gets to direct the breeding programme. Sloterdijk turns to Plato's *Statesman* for guidance. In that text, Plato first equates statesmanship to the herding of hornless, featherless land-dwelling bipeds.[165] He then compares the good ruler to a weaver, interlacing the two strands within the human character – the warlike and contemplative personalities – into an optimal social fabric.[166] ("Inappropriate" fibres are discarded.) The ruler's ultimate goal, as Sloterdijk puts it, is to produce the optimum homeostasis in the human zoo.[167] What Sloterdijk takes from Plato is the idea that more than one way of governing the breeding-programme is possible. The zookeeper-king can be of two types, democratic or tyrannical. If the ruler and the subjects are of basically the same type, or different only in degree, a democratic self-shepherding is possible. The herd can elect their own shepherd. On the other hand, if ruler and subjects are fundamentally different, democracy will not be possible. The best that can be hoped for is a zookeeper-king guided by divine insight, on the Platonic model.[168] Sloterdijk says that he inclines to the democratic kind of zookeeping, noting that divine insight is not welcome in modern politics, even species-politics. Nonetheless, he suggests that we lack guidance of any sort, whether from God or the thick letters of the

---

[164] Sloterdijk, p 25.

[165] Plato, *Statesman* (Cambridge: Cambridge University Press, 1995), 267a-c.

[166] Ibid, 311b-c.

[167] Sloterdijk, p 26.

[168] Sloterdijk, p 25.

classical humanists, leaving us with a worryingly rudderless breeding programme. On that ambiguous note, "Rules for the Human Zoo" concludes.

In the months following the Elmau conference, Sloterdijk's paper circulated among academics, before a version was published in *Die Zeit* on September 16th. The controversy gathered pace in July and August and peaked in September. Sloterdijk was excoriated. Many commentators took him to be arguing for an authoritarian programme of eugenics, aiming to "improve" the human race. Parallels were immediately drawn between his proposals and the genocidal policies of the Nazis. The journalist Thomas Assheuer said that Sloterdijk "imagines a working group unencumbered by democracy and composed of real philosophers and appropriate geneticists, who no longer debate moral questions, but rather take practical action. This élite band is entrusted with the task of initiating the genetic revision of the history of the human species by means of selection and breeding."[169] Many objected to Sloterdijk's application of words like "*Züchtung*" and "*Selektion*" to human beings.[170] Ernst Tugendhat, for one, questioned Sloterdijk's choice of "*Selektion*," selection in the sense of picking out, rather than "*Auswahl*," selection in sense of choosing. "When I listen to this word in this context," says Tugendhat, "I cannot help thinking about selection at the ramp of Auschwitz."[171]

Habermas took no part in the controversy. Nonetheless, he was one of the targets (alongside Assheuer) of an extremely vitriolic open letter from Sloterdijk, published in *Die Zeit* on September 9th. The tone of the letter can be gauged from its title: "Critical Theory is Dead."[172]

---

[169] Thomas Assheuer, "Das Zarathustra-Projekt: Der Philosoph Peter Sloterdijk fordert eine gentechnische Revision der Menschheit," *Die Zeit* 2/9/99. Quoted in Sigrid Graumann, "Experts on philosophical reflection in public discourse - the German Sloterdijk debate as an example," *Biomedical Ethics* 5:1 (2000).

[170] See Mary Varney Rorty, "For Love of the Game: Peter Sloterdijk and Public Controversies in Bioethics" (http://web.stanford.edu/~mvr2j/for_love.html, accessed 8/2/2015).

[171] Ernst Tugendhat, "Es gibt keine Gene für die Moral," *Die Zeit* 23/9/99, quoted Graumann op. cit.

[172] Peter Sloterdijk, "Die Kritische Theorie ist tot," *Die Zeit* 9/9/99.

Sloterdijk accused Habermas of working against him behind the scenes, by encouraging his academic colleagues and associates in the media, such as Assheuer, to go on the attack. He says that "Ayatollah Habermas" has issued a "Starnberg fatwa" on him. He turns the accusation of fascism back on his critics, accusing them of "liberal Jacobinism." He even charges Habermas with performative contradiction, arguing that making others speak on your behalf goes against the principles of discourse. The letter ends with Sloterdijk proclaiming that "the era of the hypermoral sons of National Socialist fathers has come to an end... Critical Theory is, on this Second of September, dead. She was long since bedridden, the sullen old woman, (and) now she has passed away completely."[173] Habermas replied with a letter to *Die Zeit* on September 16th, in which he rebuts Sloterdijk's accusations but says little about the content of "Rules for the Human Zoo."[174]

On its own, the controversy caused by Sloterdijk's paper simply demonstrates that the ethics of human genetic modification was a live issue in the last years of the 20th century – and perhaps that the people of Germany, with their collective memories of the Nazi era, were particularly sensitive to this issue. Habermas seems to have played only a minor, reluctant part in the dispute. In the years following 1999, however, he published a number of texts which deal with the issue of genetic modification, and raise the question of whether it can be addressed from a postmetaphysical standpoint.[175] The shadow of "Rules for the Human Zoo" lies most

---

[173] Ibid, pp 8-9. See also Fisher op. cit. Sloterdijk was surely aware of the effect which talk of eugenics and human breeding programmes would have in Germany, especially when combined with references to Heidegger and Nietzsche. Sloterdijk is an expert at provoking controversies, as his 2009 article "The Revolution of the Giving Hand" – which charges the welfare state with robbing its productive citizens in order to support the unproductive – demonstrates. (See Peter Sloterdijk, *Die nehmende Hand und die gebende Seite* (Suhrkamp, 2010).) In 1999, he may simply have misjudged the degree of anger his paper would provoke. The questions of Sloterdijk's personal animosity towards Habermas, and of whether there was any truth in his allegations, are outside the scope of this text.

[174] Jürgen Habermas, "Post vom bösen Geist," *Die Zeit* 16/9/99.

[175] See "Are There Postmetaphysical Answers to the Question: What is the 'Good Life'?" (FHN 1-15) "The Debate on the Ethical Self-Understanding of the Species" (FHN 16-74) and its "Postscript" (FHN 75-100), "'I Myself am Part of Nature' – Adorno on the Intrication of Reason in Nature: Reflections on the Relation between Freedom and Unavailability." (BNR 181-208)

clearly on "The Debate on the Ethical Self-Understanding of the Species," which contains

unmistakeable allusions to the Sloterdijk incident. In the introduction to that text, Habermas

says that a

> handful of freaked-out intellectuals is busy reading the tea leaves of a naturalistic version of posthumanism, only to give, at what they suppose to be a time-wall, one more spin – 'hypermodernity' against 'hypermorality' – to the all-too-familiar motives of a very German ideology… Self-styled Nietzscheans, indulging in fantasies of the 'battle between small-scale and large-scale man-breeders' as the 'fundamental conflict of all future,' and encouraging the 'main cultural factions' to 'exercise the power of selection which they have already gained,' have, so far, succeeded only in staging a media spectacle. (FHN 22)

It is clear that in this passage Habermas is referring to, and all but quoting, Sloterdijk's paper and

the controversy which followed it.

The Sloterdijk incident is significant, I suggest, because it forced Habermas to confront

the magnitude of the biotechnological threat to the human form of life, and the limits of his

ability as a postmetaphysical thinker to respond to it. Sloterdijk's talk of minimizing and

maximizing breeding programmes, of a "species politics" inspired by the ideas of Plato and

Nietzsche, showed Habermas one possible direction in which the species-ethical discourse

opened up by genetic modification might develop. It was the failure of his attempts to find a

postmetaphysical way to engage with this species-ethical discourse, I argue, that pushed him

into taking his postsecular turn in 2001. In order to make this argument, I must explain the

nature of the threat which genetic modification, according to Habermas, poses to the ethical

self-understanding of the species, and how his postmetaphysical response to it falls short. This

is what I will do in the final sections of Part 2.

2.9 Genetic Modification

"Genetic manipulation," says Habermas, "could change the self-understanding of the species in so fundamental a way that the attack on modern conceptions of law and morality might at the same time affect the inalienable normative foundations of societal integration." (FHN 26) He evidently thinks that the genetic modification of human beings, by human beings which is now becoming possible through the development of biotechnology presents a serious threat to the ethical self-understanding of the species. Let us first examine the nature of this threat, before turning to Habermas' attempts to counter it in a postmetaphysical manner.

Habermas does not oppose all genetic interventions. He distinguishes between *positive* or *liberal eugenics*, which aims to enhance human beings, and *negative* or *therapeutic eugenics*, which aims only at preventing serious harm. (FHN 19) Interventions which enhance a child's physical fitness or intelligence, or implant them with talents or dispositions desired by their parents, would count as positive eugenics. Negative eugenics, in contrast, would be limited to preventing a child from inheriting damaging hereditary conditions, through techniques such as pre-implementation genetic diagnosis. (FHN 17) Habermas objects only to positive eugenics.[176] He argues that the crucial factor in determining the acceptability of a genetic intervention is the *attitude* in which it is undertaken.

In negative eugenics, "we approach the embryo as the second person he will one day be." (FHN 43) The genetic engineer assumes, counterfactually, that the future person who they are in the process of modifying will one day be able to agree with them about the beneficial nature of the modification. It follows that only those modifications which they can be confident that a future person would agree to, such as the avoidance of debilitating diseases, may be

---

[176] Habermas' reasons for objecting to liberal eugenics are frequently misunderstood, especially by commentators who are not familiar with the broader concerns of his philosophical project. See Bernard G. Prusak, "Rethinking 'Liberal Eugenics': Reflections and Questions on Habermas on Bioethics." *Hastings Center Report* 35:6 (November-December 2005).

carried out. (FHN 63) Negative eugenic interventions are undertaken from the second person perspective, as if the genetic engineer and the engineered person were entering into intersubjective discourse. The genetic engineer certainly objectifies the embryo to some extent – they regard it as a thing to be worked on – but no more than a surgeon objectifies their patient. "This clinical attitude," says Habermas, "draws its legitimising force from the well-founded counterfactual assumption of a possible consensus reached with another person who is capable of saying yes or no." (FHN 43) The vital intersubjective relation between human beings, the cornerstone of our form of life, remains.

Positive eugenics, according to Habermas, does not consider the future agreement of the modified person. In this case the genetic engineer "approach[es] the embryo in the objectivating attitude of the technician, that is, as an object which is manufactured or repaired or channelled into a desired direction." (FHN 52) The embryo is regarded from the third person perspective, as raw material on which to imprint the desires of parents or designers. There is no clinical attitude, no intersubjective relation here. Positive eugenics is closer to manufacturing than healing, and the objectivating attitude under which it is undertaken turns the embryo from a quasi-subject into an object. The implications, for Habermas, are profound. "What is at stake, therefore, with the instrumentalization of prepersonal life is the ethical self-understanding of the species, which is crucial for whether or not we may go on to see ourselves as beings committed to moral judgement and action." (FHN 71)

So Habermas believes that positive eugenics encourages human beings to see each other in subject/object rather than intersubjective terms. The same could be said, of course, about racism, sexism and similar phenomena. Habermas does not consider these distortions of intersubjectivity to endanger the ethical self-understanding of the species. What is special about genetic modification?

As we have seen, Habermas sees morality as grounded in the intersubjective world. It is by rationally reconstructing the communicative processes of the intersubjective world, making explicit their implicit normativity, that he derives his discourse theory of morality. For morality to be universal, all subjects must, in principle, be able to participate in unforced, symmetrical communicative action with each other.[177] Of course, this assumption is often fulfilled only counterfactually. In societies as divided and unequal as our own, the preconditions for unforced symmetrical discourse are often missing. Nonetheless, Habermas maintains that it has always been possible *in principle* for any two contemporary human beings to engage in communicative action. This is what makes discursive morality universal. "The conviction that all actors, as persons, obtain the same normative status and are held to deal with one another in mutual and symmetrical recognition," Habermas asserts, "rests on the assumption that there is, in principle, a reversibility to interpersonal relationships." (FHN 63)

Genetic modification destroys this implicit symmetry. Liberal eugenics creates "an interpersonal relationship for which there is no precedent," (FHN 63) a relationship which is *inherently asymmetrical.* Unlike the contingent and reversible relations of dependency which exist between members of unequal societies, the ontological dependency of the designed person on their designer is *irreversible*.[178] "The product cannot, to put it bluntly, draw up a

---

[177] There are of course limits to Habermas' approach. As he himself acknowledges, subjects who are beyond the range of communication, such as the dead and non-human animals, cannot be regarded from the moral point of view. Questions of past injustices and the treatment of non-human life thus occupy a troubling position for the discourse theory of morality. In "Morality and Ethical Life," written around 1986, Habermas accepted that these are "problems that flow from the self-limitation of every nonmetaphysical point of view," (MCCA 210) and that "Since the concept of morality is limited, the self-perception of moral theory should be correspondingly modest." (MCCA 211)

[178] The closest thing to an inherently asymmetrical relationship, outside the realm of liberal eugenics, is the parent/child relationship. After all, parents raise children, and not vice versa; the influence that parents exert on the development of their children does not flow in the opposite direction. However, Habermas argues that there is a qualitative difference between parental relationships and eugenic ones. Since it is formed communicatively, in the intersubjective world, the relationship of parental dependence is reversible. Adolescents can "critically reappraise" and contest the traces of their upbringing; someone raised with the expectation that they will be a musical prodigy, for example, can choose a different course in life. By doing so, they establish symmetrical relations with their parents. (FHN 61-2) "Even neurotic fixations," says Habermas, "may be resolved analytically, through an elaboration of self-reflexive insights." (FHN 62) This is not possible in cases of genetic modification. (FHN

design for its designer." (FHN 65) It would not be possible, according to Habermas, for symmetrical intersubjective relations to exist between designer and designed. As far as he is concerned, this is something completely new in human history. For the first time, two contemporary human beings would not be able to enter into symmetrical intersubjective relations with each other, or to engage in communicative action. If we grasp the importance of this for Habermas, then the part played by genetic modification in pushing him towards the postsecular turn becomes easier to understand. He does not regard genetic modification as simply a morally problematic development – he regards it as an anthropic catastrophe. Genetic modification breaks up the intersubjectivity which has always been part of the human form of life, replacing it with instrumental subject/object relations *which cannot be overcome in discourse.* This is what Habermas means when he says that biotechnology undermines not only law and morality, but "the inalienable normative foundations of societal integration." The species-wide phenomena which he rationally reconstructs in his theories of formal pragmatics, morality, and law and democracy would no longer be species wide; the intersubjective basis for his philosophical project would fall to pieces.

As with hard naturalism, Habermas needs to find a way of arguing against the genetic modification of human beings, by human beings. Once again, however, the anthropic problem obstructs him. The question of whether the human form of life should be based on intersubjective or subject/object relations is a question of the ethical self-understanding of the species. As a postmetaphysical philosopher, Habermas has disbarred himself from taking part in species-ethical discourse. This case is, if anything, more serious than the last. Hard naturalism is a theory, which Habermas believes to be based on faulty logic. He can show that the theoretical assumptions of hard naturalism are incoherent, and thus counter the theory using a

---

51) A child who has been *programmed*, rather than socialized, with musical talents has nothing to critically reappropriate; their only choices are submission or alienation. "*Being at odds with* the genetically fixed intentions of a third person is hopeless." (FHN 62)

postmetaphysical argument. Genetic modification is a *practice*, which is already taking place and looks certain to become more widespread in the future. It is not a theory that Habermas can argue against. The only way to counteract it is to make the case that the human race's current form of life is preferable to one structured by biotechnological interventions. Sloterdijk's "Rules for the Human Zoo" showed Habermas the territory into which this debate could move, at its worst. The necessity of participating in the species-ethical discourse around genetic modification, then, is clear. But, constrained as he is by the postmetaphysical paradigm, what can Habermas do?

In his search for a way out of this conundrum, Habermas pushes his paradigm to its limit. In the short text "Are there Postmetaphysical Answers to the Question: What is the Good Life?," he attempts to construct a *postmetaphysical ethical theory*, which will allow him to intervene in species-ethical discourses. This is a peculiar text, in which Habermas seems to be pulled in two directions. On the one hand, he is aware that for the sake of being consistent with his longstanding philosophical commitments, his answer to the title question should be "no." He restates his belief in the priority of morality over ethics, and the ethical abstinence of postmetaphysical thinking, saying that "today, in our postmetaphysical age, philosophy no longer pretends to have answers to questions regarding the personal, or even the collective, conduct of life," (FHN 1) and that practical philosophy must restrict itself to "questions of justice. In particular, its aim is to clarify the moral point of view from which we judge norms and actions whenever we must determine what lies in the equal interest of everyone and what is equally good for all." (FHN 3) On the other hand, Habermas appreciates how urgent the need for species-ethical discourse has become: "This postmetaphysical abstention runs up against its limits in an interesting way as soon as questions of a "species ethics" arise. As soon as the ethical self-understanding of language-using agents is at stake *in its entirety,* philosophy can no longer avoid its substantive position (…) It is just this situation that we find ourselves in today," due to

the "advance of the biological sciences and development of biotechnologies." (FHN 11) Postmetaphysical philosophy should not become entangled in ethical discourse; and yet it must.

Ethical abstinence is a tenet of postmetaphysical thinking. Nonetheless, Habermas thinks that he can find a postmetaphysical point of entry into ethics. "Kierkegaard," he says, "was the first philosopher who answered the basic ethical question regarding the success or failure of one's own life with a postmetaphysical concept of 'being-able-to-be-oneself'… Kierkegaard answered the question of the right way to live with an answer that was indeed *postmetaphysical*, while at the same time *theological*." (FHN 5)[179] Habermas believes that several of Kierkegaard's central concepts – the ethical stage of life, the "sickness unto death," and the dependence of the subject on a "wholly Other" – can be worked up into an ethical theory which meets the requirements of postmetaphysical thinking.

A person at Kierkegaard's ethical stage is aware of their own freedom and capable of taking responsibility for their actions.[180] Habermas sees this as an account of successful subjectivity, of what is required for a subject to play the language game of responsible agency. Kierkegaard's attention, he says,

> is on the structure of the ability to be oneself, that is, on the form of an ethical self-reflection and self-choice that is determined by the infinite interest in the success of one's own life-project. With a view toward future possibilities of action, the

---

[179] We saw in § 2.3 that Habermas is inconsistent in his classification of Kierkegaard. He often regards him as a postmetaphysical thinker, but in contrasting both Fichte and Kierkegaard's theories of subjectivity to that of Mead (PT 24-6, 161), he seems to place him in the subjectivist category. In this passage that I have quoted Habermas adds that Kierkegaard's thought is simultaneously postmetaphysical and theological, in the sense that he does not adhere to methodological atheism, like later existentialists such as Heidegger, Jaspers and Sartre. (FHN 5-6) Habermas' classification of Kierkegaard is debatable. There are many grounds for regarding him as a metaphysical or subjectivists thinker. Kierkegaard's entire philosophical project could be described as ethical-existential discourse. He continually uses metaphysically-charged concepts from the Judaeo-Christian tradition (God, sin, forgiveness, redemption, the story of Abraham and Isaac), and if these are conceptual appropriations, they have not passed through the filter of methodological atheism. Many of his central concepts (the teleology of the aesthetic, ethical and religious stages, the dependence of the subject on a wholly Other) appear to be prototypically metaphysical.

[180] Søren Kierkegaard, *Either/Or: A Fragment of Life*, vols. 1 and 2 (London: Oxford University Press, 1946).

individual self-critically appropriates the past of her factually given, concretely re-presented life history. Only then does she make herself into a person who speaks for herself, an irreplaceable individual. (FHN 6-7)

The ability to critically appropriate one's past is one of the requirements of successful subjectivity. The antithesis of the ethical stage – or the prompt to go beyond it – is the syndrome of conditions which Kierkegaard calls the "sickness unto death."[181] "These forms of despair," says Habermas, "are so many manifestations of the lack of a fundamental relationship that alone could make an authentic being-oneself possible." (FHN 8) Sickness unto death is in effect an inability or refusal to be a self, which Habermas connects to the inability or refusal to critically appropriate one's life history. Kierkegaard thinks that this failure of subjectivity forces the subject to realise that (contra Fichte) the ground of its freedom is not in itself, but in a "wholly Other," something outside the subject which transcends it. For Kierkegaard, this Other is the Christian God. It is only possible to successfully be oneself in relation to God. Of course, this is not acceptable by the standards of postmetaphysical thinking. A secular equivalent to the concept of "god" is needed.[182] Habermas finds this in language and communication: "The linguistic turn permits a deflationary interpretation of the 'wholly Other.' As historical and social beings we find ourselves always already in a linguistically structured lifeworld... The *logos* of language embodies the power of the intersubjective, which precedes and grounds the subjectivity of speakers." (FHN 10-1) To be able to be myself, then, involves acknowledging the dependence of my subjectivity on a prior, linguistically structured intersubjectivity.

Habermas re-works Kierkegaard's thought, or perhaps conceptually appropriates some of its motifs. He believes that this yields an ethical theory, centred on the "ability to be a self," (*Selbstseinkönnen*) which meets the standards of postmetaphysical thinking. The theory makes no metaphysical claims, according to Habermas; the *logos* of language is not extramundane, it

---

[181] Søren Kierkegaard, *The Sickness Unto Death: A Christian Psychological Exposition for Upbuilding and Awakening* (Princeton, New Jersey: Princeton University Press, 1980).

[182] Why, then, does Habermas insist on classifying Kierkegaard as a postmetaphysical thinker?

is situated within our lifeworld practices, and the "unconditionedness of truth and freedom is a necessary presupposition of our practices, but beyond the constituents of 'our' form of life they lack any ontological guarantee." (FHN 11) The theory's "general statements about the modes of being-able-to-be-oneself are formal – that is, they are not *thick* descriptions – but they by no means lack normative content. Because this ethics judges the existential *mode*, but not the specific orientation of, individual life-projects and particular forms of life, it satisfies the condition of a pluralism of worldviews." (FHN 11) As Junker-Kenny puts it, "In contrast to approaches that identify specific goods and virtues to be pursued in a life, '*Selbstseinkönnen*' is a formal qualifier for every life project. Kierkegaard's ethics establishes a foundational level that does not undermine modernity's distinction between the good and the right."[183] To Habermas' satisfaction, at least, this means that his adaptation of Kierkegaard's theory is both ethical and postmetaphysical. His next move is to apply the theory to the species-ethical discourse around genetic modification. If he is to argue against genetic modification, he must show that it impairs a person's ability to successfully be a self.

Habermas does this by identifying some of the biological preconditions for being a self, and arguing that genetic modification disrupts them. "The capacity of being oneself requires that the person be at home, so to speak, in her own body." (FHN 57) Making use of Arendt's concept of natality (FHN 58-9) and Adorno's idea that freedom is connected to *unavailability* (*Unverfügbarkeit*) (BNR 201), Habermas argues that to function as subjects, we must conceive of some aspects of our nature as being unavailable. For me to understand myself as free, I must know that no-one, myself included, has disposed over my inner nature or biological origins:

> We experience our own freedom with reference to something which, by its very nature, is not at our disposal. The person, irrespective of her finiteness, knows herself to be the irreducible origin of her own actions and aspirations. But in order to know this, is it really necessary for this person to be able to ascribe her own origin to a beginning which eludes human disposal, to a beginning, that is, which is

---

[183] Junker-Kenny, *Habermas and Theology*, p 124.

sure not to prejudge her freedom only if it may be seen as something – like God or nature – that is not at the disposal of some *other* person? (FHN 58)

To understand myself as created by God or nature, or to view my subjectivity as a product of a prior intersubjectivity, is not harmful to my ability to be a self, since the same is true for everyone. We are all "persons of equal birth" in this sense. (FHN 64) Being created, like any other human being, by conditions outside my control leaves my subjectivity intact. Being created by another human being is a different matter. When a genetic programmer, says Habermas, "makes himself the *co-author of the life of another*, he intrudes – from the interior, one could say – into the other's consciousness of her own autonomy. The programmed person, being no longer certain about the contingency of the natural roots of her life history, may feel the lack of a mental precondition for coping with the moral expectation to take, even if only in retrospect, the *sole* responsibility for her own life." (FHN 81-2) Removing my certainty that my origins were not disposed over may remove my ability to appropriate my own life history, and therefore to be a self.

Having established this association between having "unavailable" origins and being able to be a self, Habermas turns to the distinction between the *grown* and the *made*, or the natural and the artificial. Junker-Kenny says that he uses "Kierkegaard's concept of 'being able to be oneself'… as the critical standard from which 'the grown and the made' can be distinguished as the un-manipulated and the imposed."[184] Human beings have long made a phenomenological distinction between the grown and the made. That which is made is at our disposal, and that which is grown is not. Although we have hitherto understood ourselves as grown, advances in biotechnology are now starting to blur this longstanding categorisation. Genetic modification of human beings, by human beings,

> uproot(s) the categorical distinction between the subjective and the objective, the naturally grown and the made, as they extend to regions which, up to now, we could not dispose over. What is at stake is a dedifferentiation, through

---

[184] Junker-Kenny, *Habermas and Theology*, p 119.

biotechnology, of deep-rooted categorical distinctions which we have as yet, in the description we give of ourselves, assumed to be invariant. (FHN 42)

The reason this matters, says Habermas, is that my understanding of myself as grown, like my knowledge that my origins were not disposed over by another, may be necessary for me to successfully be a self, and for all that follows from this:

This dedifferentiation might change our ethical self-understanding as a species in a way that could also affect our moral consciousness – the conditions, that is, of nature-like growth which allow us to conceive of ourselves as the authors of our own lives and as equal members of the moral community. (FHN 42)

If I understand myself as made rather than grown, I am aware that another person has disposed over my origins and composition. Habermas uses his Kierkegaardian ethical theory to suggest that this awareness may block my ability to see myself as the author of my own life. Without the capacity for selfhood, Habermas argues, human life would be drastically different, and much worse than it is with our current ethical self-understanding. A human being who was not able to be a self would be unable to take moral responsibility for "their" own actions (FHN 57), and thus unable to be part of the moral community. They would view themselves from the third person instead of the first person perspective. (FHN 53) They would *have* a body, but not *be* a body. (FHN 54)[185]

This is how Habermas uses his postmetaphysical ethical theory, derived from Kierkegaard, to intervene in the species-ethical discourse opened up by genetic modification. The upshot of his arguments is that if we want to retain our ability to be subjects, we must retain the ethical self-understanding which grounds it. This involves having "unavailable" origins, seeing ourselves as grown rather than made, and so on. Since liberal eugenics makes our origins "available" and forces us to see ourselves as made rather than grown, we should reject it. The question is, are Habermas' arguments against genetic modification successful?

---

[185] Habermas is referring to Helmuth Plessner's distinction between being a body and having a body. See FHN 50.

There are several reason for concluding that they are not. To begin with, Habermas' diagnosis of the dangers of genetic modification is highly speculative. He suggests that genetic modification would disrupt subjectivity, shift the ethical self-understanding of a modified person in the direction of being made rather than grown, and so on. In reality, biotechnology is in its infancy, and we simply cannot know if it would have any detrimental effect on the ability to be a self. Habermas is aware of this. "Who knows," he asks, "whether knowledge of the fact that the makeup of my genome was designed by someone else need be of any significance at all for my life? It is rather unlikely that the perspective of being a body will lose its primacy over having a genetically tailored body… Why should people not get used to this, too, and shrug it off by saying 'so what?'?" (FHN 54) After all, human beings have come to terms with the knowledge that we are naturally evolved rather than divinely created. It may be that subjectivity is more flexible than Habermas imagines. In the absence of evidence that genetic modification would impair the modified person's ability to be a self, he cannot rule the practice out a priori; the most he can do is to urge caution. Of course, such evidence might come to light after the fact, but this would be too late for Habermas' purposes. By the time it became clear that genetic modification was a threat to subjectivity, the damage to the human form of life which he fears would already have occurred.

More seriously, Habermas' Kierkegaardian theory cannot play the role he wants it to play in the species-ethical discourse surrounding genetic modification. The theory is hampered by his attempt to make it both ethical and postmetaphysical. Habermas claims that that the theory is postmetaphysical in two ways. Firstly, its account of the ground of subjectivity deflates Kierkegaard's transcendent, metaphysical concept of God into the logos of language, which is immanent to the human lifeworld. Secondly, it gives a formal description of the conditions for being a subject able to pursue the good life, rather than a substantive account of the good. (FHN 11) It is this second point which undermines Habermas' attempt. He can do a great deal of *diagnostic* work with his postmetaphysical ethical theory. He can argue that human subjectivity

has hitherto been of a certain type, and explain that it grounds our current morality and form of social integration. He can give a detailed account of how genetic modification might affect our ability to be selves, and thus our form of life. But what he cannot do is say anything *prescriptive*, about which form of life is best for us as a species. A description of the human life-form and of how biotechnology might change it, however sophisticated, is not an argument for retaining that life-form.

Habermas is intermittently aware of this problem. "It cannot be taken for granted," he admits, "that we will still w*ant* this status of a member of a community that requires all its members to show equal respect for every other member and to be responsible in their solidarity with all of them… why – if biotechnology is subtly undermining our identity as members of the species – should we *want* to be moral?" (FHN 73) Granted that genetic modification is changing our form of life, which grounds our status as moral subjects; why should we retain this form of life? Habermas knows that his postmetaphysical ethical theory cannot answer these questions. He carefully ensures that it contains a formal description of the "modes of being-able-to-be-oneself," rather than "*thick* descriptions" of the good. (FHN 11) But thick descriptions of the good, substantive ethical values, are precisely what Habermas needs if he is to argue that one form of life is better for the human species than another. In the "Postscript" to "The Debate on the Ethical Self-Understanding of the Species," he says that "there is one argument that acquires particular weight in the debate over the best ethical self-understanding of our species: not all of the ethical conceptions harmonize with our self-understanding as morally responsible persons to the same degree." (FHN 93) This is no argument at all, since it assumes that retaining our current self-understanding as morally responsible persons is a priori desirable. This is the very point which the species-ethical discourse opened up by genetic modification has called into question. Confronted by the demand for reasons for preferring one form of life over another – for species-ethical values – Habermas' theory falls silent. He concludes "The Debate on the Ethical Self-Understanding of the Species" by saying that life without moral sentiments and a

solidary lifeworld would be unbearable. (FHN 73) It certainly would *for beings like us*. Future human beings, however, may not feel the same way.

The fundamental problem with any postmetaphysical attempt to intervene in species-ethical discourse is that postmetaphysical theories are rationally reconstructions *of a form of life*. This is the inescapable consequence of Habermas' programmatic decision to make his philosophical project immanent rather than transcendent, postmetaphysical rather than metaphysical. Since they are immanent to a particular form of life in this way, postmetaphysical theories cannot yield reasons for preferring one form of life over another. The species-ethical values which Habermas needs if he is to defend our current form of life from biotechnological alterations *cannot come from within his own paradigm*. In the case of genetic modification, Habermas does not succeed in arguing for the retention of the current ethical self-understanding of the species.

2.10 Postmetaphysical Thinking and the Postsecular Turn

Habermas' paradigm has an unstated species-ethical basis, and at the same time purports to be ethically abstinent. This becomes a problem when new ideas or practices threaten to change the ethical self-understanding of the human species. Habermas, as a postmetaphysical philosopher, cannot intervene in the species-ethical discourses which could undermine his entire philosophical project. This is what I call the anthropic problem. By the end of the 20th century, Habermas was aware of two such threats to the ethical self-understanding of the species: hard naturalism and genetic modification. He succeeds in making a postmetaphysical argument against the former; he does not succeed with the latter.

The problem with Habermas' attempt to argue against genetic modification in a postmetaphysical manner is that his paradigm cannot supply, from its own resources, the substantive ethical values which a species-ethical discourse calls for. Ethical values, as far as he is concerned, are bound up with metaphysical claims about reality, the good for human beings, and the teleology of life and history; moreover, they are relative to particular individuals and communities, and thus not suited to the postmetaphysical method of rational reconstruction. Short of abandoning his entire approach to philosophy, Habermas cannot supply the ethical values which he needs in order to argue against genetic modification. Faced with the species-ethical threat of biotechnology, postmetaphysical thinking has nothing to offer.

There is, of course, one way out of this dead end for Habermas – he could borrow substantive ethical values form another source, and use them to argue that the current human form of life is better for us than any other. As we have seen, there are precedents for such borrowings in his body of work. In section 2.5, I described Habermas' procedure of ethical appropriation, which makes good the ethical abstinence of postmetaphysical thinking. In section 2.6, I gave an account of Habermas' approach to conceptual appropriation, a procedure which many philosophers have engaged in. There are hints throughout "The Debate on the Ethical Self-

Understanding of the Species" that Habermas has a particular source of ethical values in mind. Early in the text, he says that modern lifeworlds have to rely on "their own secular resources for regenerating the energies that ensure their own moral cohesion." He worries about the "rigidity" of secular lifeworlds, "which, having lost their backing of metasocial guarantees, are no longer able to respond to new threats to their sociomoral cohesion by new secularizing impulses, let alone by yet another moral and cognitive recasting of religious traditions." (FHN 26) Later, Habermas says that religious and metaphysical traditions express "an anthropological self-understanding of the species that is consistent with an autonomous morality. The religious interpretations of the self and the world that were elaborated by highly advanced civilizations during the axial age converge, so to speak, in a minimal ethical self-understanding of the species sustaining this kind of morality." (FHN 40) Religious traditions, certainly, are sources of powerful ethical values. By using Kierkegaard's thought as his point of entry into the species-ethical discourse surrounding genetic modification, Habermas had already come close to borrowing religious reasons and arguments.

Habermas' text "Are There Postmetaphysical Answers to the Question: What is the 'Good Life'?," in which he outlines his supposedly postmetaphysical ethical theory, was delivered as a lecture in September 2000. "The Debate on the Ethical Self-Understanding of the Species," which he describes as "an entrance into this debate that does not relinquish the premises of postmetaphysical thinking," (FHN vii) was presented at the end of June 2001. In September, Islamic extremists attacked New York and Washington DC, killing three thousand people. A month after that, Habermas took to a podium in the Paulskirche in Frankfurt, to deliver a paper entitled "Faith and Knowledge." In this text, he first discusses the idea of postsecularism.

A number of factors pushed Habermas towards making his postsecular turn. As we will see in Part 3, he claims that his new model of postsecular deliberative democracy is meant to alleviate the asymmetrical cognitive burden which the secular public sphere places on religious

citizens, in the interests of political legitimacy as much as equity. These moral and political considerations undoubtedly played a part in his change of direction. In my view, however, the endogenous factor of the anthropic problem, which few commentators have explored, was decisive. Habermas had realised, by 2001, that his postmetaphysical paradigm was incapable of defending its own species-ethical basis, and indeed incapable of making any arguments which could defend the human form of life from the potentially horrifying consequences of genetic modification. Adopting postsecularism gave Habermas a *deus ex machina* solution to this problem. On the one hand, he could continue to pursue his postmetaphysical research programmes. On the other, he could import ethical values, reasons, and arguments from religious sources into the public sphere, in order to take part in the species-ethical discourse around genetic modification.

At the centre of postsecular deliberative democracy, as we will see, is what Habermas calls the "translation" of religious inputs to the public sphere into secular language. I will have a great deal to say about this procedure in Part 3. Habermas claims that the function of translation is to balance the cognitive burdens borne by religious and non-religious citizens, and to ensure that religious believers can take part in the legitimising process of public discourse. In practice, however, he seems to regard the procedure as a means of accessing the species-ethical values embedded in religious traditions. As Darren Walhof puts it, "the purpose of Habermas' translation requirement appears to be less about ensuring that religious citizens have full religious liberty and can recognize themselves as authors of the laws under which they live, and more about the mining of theological resources for public benefit."[186]

Habermas' discussions of the translation procedure make this clear. In all his writings on postsecular deliberative democracy, he only gives one example of translation, and this example concerns "how we should deal with prepersonal human life under descriptions of molecular

---

[186] Darren R. Walhof, "Habermas, same-sex marriage and the problem of religion in public life" *Philosophy and Social Criticism* 39 (March 2013), p 235.

biology that make genetic interventions possible." (FHN 108) In "Faith and Knowledge," he says that Christian groups in Germany have translated the religious concept of man being created in the image of God ("So God created man in his own image, in the image of God created he him") into the "secular language of the constitution" as the statement that "the gamete fertilized ex utreo" is "a subject of human rights." (FHN 109) Habermas goes on to say that this secular translation of a religious concept can illuminate the difference between the dependence of created beings on God, and the "entirely different kind of dependence" which exists if "the place of God is taken by a peer – if, that is, a human being would intervene, according to his own preferences and without being justified in assuming, at least counterfactually, a consent of the concerned other, in the random combination of the parents' sets of chromosomes." (FHN 115) In other words, Habermas' sole example of postsecular deliberative democracy in action involves the species-ethical discourse around genetic modification. Here ethical values "translated" from religious into secular language are used to reaffirm the value of the current ethical-self-understanding of the species, and to argue against any biotechnological changes to it.

Religion, in translation, plays the role that postmetaphysical philosophy, trapped by the anthropic problem, could not. Habermas' turn to postsecularism in 2001 was motivated by his paradigm's inability to tackle the urgent species-ethical questions raised by genetic modification. In the final part of this text, I will turn to his model of postsecular deliberative democracy. We have examined the reasons for the postsecular turn; it is now time to evaluate the result.

Part 3

Translating the Sacred

The heavens are the Lord's heavens – but the Earth He has given to the sons of men.

Psalm 115

That would be very nice, but it is hardly translation.

Ezra Pound, "Guido's Relations"

3.1 Introduction – The Postsecular Turn

On October 14th 2001, at the Paulskirche in Frankfurt, Habermas delivered a speech on receiving the Peace Prize of the German Booksellers and Publishers Association. It is in this text, "Faith and Knowledge," that he first uses the term "postsecular," and starts to modify the discourse theory of law and democracy, so as to allow religion a place in the public sphere. I will refer to Habermas' new political model as *postsecular deliberative democracy*. He expands on it in texts such as 2005's "Religion in the Public Sphere," 2008's "What is Meant by a "Post-Secular Society"?," and others. At the time of writing Habermas is at work on a book, *Essay on Faith and Knowledge*, intended to be his definitive statement on religion.

[1] In the third part of this text I will focus on the political and moral side of his postsecularism, on the question of religion in the public sphere. I begin by recounting Habermas' justifications for adopting his model of postsecular deliberative democracy, which he defines in opposition to Rawls, (3.2-4) and the problems he identifies in his earlier model of secular deliberative

---

[1] See the numerous references in *Habermas and Religion,* op. cit., and especially the appendix by Eduardo Mendieta, "Religion in Habermas' Work," pp 707-709.

democracy (3.5-6). Objections are raised to some of Habermas' arguments on these points. Next I will describe Habermas' new model (3.7-8) and criticise its central procedure, *sacred-to-secular translation* (3.9-10). I conclude that the problems of the translation procedure makes Habermas' model unacceptable.

3.2 Justifying the Postsecular Turn

Habermas offers three justifications, often overlapping, for adopting his model of postsecular deliberative democracy.[2] On the philosophical level, he wants to make the contents of religious traditions available for the ethical discourses of modern subjects. He believes that the vocabulary of religion contains valuable insights, absent from secular traditions, which can enrich modern ethical discourse. (FHN 109-11) I argued in Part 2 that Habermas is motivated here by the need to find inputs to the species-ethical discourse opened up by hard naturalism and genetic modification. In my view, this is the most important factor behind the postsecular turn; but since I have already discussed it extensively, I will put it aside in what follows.

On the political level, Habermas wants to protect some variety of secular, liberal, democracy from the violent reaction which confronts it. "Faith and Knowledge," written in October 2001, links the disaffection which secular democracy supposedly[3] provokes in religious believers to the recent surge of terrorist violence. The accelerated process of secularization which economic development forces on the non-Western world, argues Habermas, leads to anomie, resentment and a compensatory turn to fundamentalism.[4] (FHN 102) To combat these trends, we should no longer assume that secularism of the European type is the end-point of

---

[2] Maeve Cooke describes three motivations for the postsecular turn which are similar to the ones I list here: the ethical issues raised by biotechnology, the rise of violent fundamentalist movements, and the moral issue of the cognitive burdens of secularism, touched on in Habermas' dispute with Rawls during the 1990s. See Cooke, "Salvaging and Secularizing the Semantic Contents of Religion," p 189.

[3] María Herrera Lima points out that there is little sociological evidence for such an attitude amongst the religious citizens of liberal democracies. See María Herrera Lima, "The Anxiety of Contingency: Religion in a Secular Age," *Habermas and Religion* p 62.

[4] Vivienne Boon argues that Habermas' account of the causes of religious fundamentalism is disappointingly reductive, coming close to a crude-Marxist economic determinism. He assumes that uneven economic development and rationalization have led to systematically distorted communication and violence; the solution is to open channels of free communication with the non-Western world. Habermas explains away fundamentalist thinking, argues Boon, without seriously engaging with it. He does not consider the possibility that some religious fundamentalist thinkers, such as Sayyid Qutb, have a basically non-discursive perspective, rendering the panacea of communicative action ineffective. See Vivienne Boon, "Jürgen Habermas and Islamic Fundamentalism: on the limits of discourse ethics," *Journal of Global Ethics* 6:2 (August 2010).

social development, and that other societies can or should follow the European path. It is the West, rather, that needs to reassess its assumptions about religion, public reason, and the state. (FHN 102-3) A strand of social-scientific postsecularism evidently underlies this argument for normative-political postsecularism: if religion is not, as Weber predicted, doomed to decline, then deliberative democracy must find a place for it. (BNR 116)

Habermas soon moves away from this these global considerations, however, and concentrates solely on Western societies. His postsecular model, he says, can apply only to societies which have experienced a period of secularism. He names the European countries (with the exceptions of Ireland and Poland)[5], Canada, Australia, and New Zealand as examples. (EFP 59) Presumably Habermas sees himself as taking part in the reappraisal of the secular West, rather than constructing a model of postsecular deliberative democracy which is applicable everywhere. Nonetheless, this narrow focus has implications for his model. All of Habermas' arguments for postsecularism assume a highly economically developed society, with a deep tradition of liberal democracy and a secular majority. He pays no heed to the two major examples of non-Western secularism, India and Turkey; he devotes little attention to the United States. Many questions, as a result, are overlooked. Habermas' model does not tell us what variety of postsecularism is appropriate for highly religious societies, or for societies with a small number of large, competing denominations, or for developing societies with little experience of formal secularism. Habermas sees the postsecular model as a way to integrate excluded religious minorities into the democratic processes of, by assumption, non-religious societies. He

---

[5] Poland, it is true, is a particularly devout country, especially compared to its immediate neighbours. But Habermas overlooks the sharp drop in both professed religious belief and church attendance observed in the Republic of Ireland in recent decades. A 2012 WIN-Gallup International poll found that the religiosity of the Irish population declined from 69% in 2005 to 47% in 2012 – the second-largest decline in the world during this period – while the percentage of the population describing themselves as "not religious" increased from 25% to 44%. See "WIN-Gallup International, Global Index of Religiosity and Atheism 2012 – Press Release," at http://redcresearch.ie/wp-content/uploads/2012/08/RED-C-press-release-Religion-and-Atheism-25-7-12.pdf (accessed 25/2/2014).

does not consider the effect his model might have in a society with a religious majority and non-religious or differently-religious minorities. [6]

The philosophical and political justifications for the postsecular model are connected. The enemy, as always for Habermas, is instrumental reason, whether it comes in the form of a naturalistic self-understanding grounded in scientific discoveries, or the destructive devaluation of the lifeworld driven by global capitalism. In both cases, he sees the incorporation of religion into the public sphere as a way of halting these process of colonization.

Habermas' third justification is moral, and looks to the failure of other political theories to deal adequately with the presence of religious belief in modern societies. Theories of secular deliberative democracy, Habermas argues, place unacceptable strains on religious citizens; they must be modified to remove this unfair burden. The theory which Habermas has in mind is John Rawls' political liberalism,[7] which he takes as a foil for his postsecular model, along with Robert Audi's arguments for strict secularism.[8] (BNR 121-4, 126-7) Habermas certainly does not want to abandon secularism; he distances his model from the anti-secular theories of Nicholas Wolterstorff[9] and Paul Weithman.[10] (BNR 123-4) As ever, he constructs an intermediate model,

---

[6] See Nadia Urbinati, "Laïcité in Reverse: Mono-Religious Democracies and the Issue of Religion in the Public Sphere" *Constellations* 17:1 (March 2010). Writing about Italy, Urbinati says that both Rawls' and Habermas' models "are tailored to a philosophical reflection of the liberal societies that are the home of genuine religious pluralism. However, they are – Habermas' more than Rawls' – hardly suitable and safe if extended to liberal societies in which one religion enjoys a strong majority and pluralism is only predicated in the constitution but is not a lived experience in society." (pp 4-5)

[7] John Rawls, *Political Liberalism* (New York: Columbia University Press, 1996), hereafter PL.

[8] See Robert Audi, "The Separation of Church and State and the Obligations of Citizenship," *Philosophy and Public Affairs* 18:3 (1989), and Robert Audi, "The Place of Religious Argument in a Free and Democratic Society," *San Diego Law Review* 30:4 (1993).

[9] See Robert Audi and Nicholas Wolterstorff, *Religion in the Public Square* (Lanham, Md.: Rowman and Littlefield, 1997).

[10] Paul J. Weithman, *Religion and the Obligations of Citizenship* (Cambridge: Cambridge University Press, 2004). Cooke has recently moved in a similar direction. See her papers "Violating Neutrality? Religious Validity Claims and Democratic Legitimacy," *Habermas and Religion* pp 249-274, and "The Limits of Learning: Habermas' Social Theory and Religion" *European Journal of Philosophy*, forthcoming.

which requires both a modification of the political institutions he described in *Between Facts and Norms* and a change in attitudes on the part of citizens. Before evaluating Habermas' moral arguments for the postsecular model, it will be necessary to review Rawls' (apparently too secular) model. Political liberalism is itself a revision of Rawls' early work, carried out in response to the existence of diverse beliefs in modern societies.[11] The problematic relationship between deliberative democracy and belief pluralism frames this entire discussion, as the next section explains.

---

[11] This is how I interpret Rawls' later work. My interpretation is in accord with the "alternate blueprint" sketched by Anthony Simon Laden in "The House That Jack Built: Thirty Years of Reading Rawls," *Ethics* 113:2 (January 2003).

3.3 The Problem – Reasonable Pluralism and Deliberative Democracy

Many theorists have identified a tension between the principles of *deliberative democracy* and the fact that modern societies contain diverse systems of beliefs. I am using "deliberative democracy" as a broad label which can be attached to any theory of democracy which grounds the legitimation of laws and policies in the public use of reason. Habermas and Rawls both have theories of this sort.[12] The basic assumption of this approach is that *agreement generated in public discourse is the key to legitimation* – only the public exchange between citizens of mutually-understandable reasons and arguments can justify coercive laws and policies. Belief pluralism puts this into question. Cristina Lafont points out that:

> only if [citizens] provide the arguments and counter-arguments they sincerely believe are right regarding the policies under discussion will they then be able to follow the 'unforced force of the better argument', to use Habermas' term, and reach a conclusion in good faith on the acceptability of those policies. However, allowing citizens to provide reasons and justifications on the basis of cognitive stances that are not shared seems directly incompatible with the democratic obligation of providing generally acceptable reasons to justify coercive policies with which all citizens must comply.[13]

Lafont characterises the problem as a conflict between citizens' cognitive obligation to examine laws and policies in the light of all the reasons they, personally, consider relevant, and their democratic obligation to provide reasons acceptable to other citizens.[14]

On the one hand, theories of deliberative democracy assume that legitimate laws and political decisions are those which all affected parties can agree to, under ideal conditions of discourse. This is the insight captured in Habermas' discourse theory of law and democracy, in the formula that citizens should ideally be able to see themselves as both the authors and

---

[12] Rawls identifies his theory as a theory of deliberative democracy in "The Idea of Public Reason Revisited," in *The Law of Peoples* (Cambridge, Mass. and London: Harvard University Press, 2001), pp 138-40. Hereafter IPPR.

[13] Cristina Lafont, "Religion and the Public Sphere: What are the Deliberative Obligations of Democratic Citizenship?" *Philosophy and Social Criticism* 35:1 (2009), pp 128-9.

[14] Ibid, p 129.

addressees of legitimate laws. (BFN 120) For such a legitimating discourse to take place, there must be a public "language," a shared set of reasons and arguments which all citizens can understand. Belief pluralism, however, means that citizens may not all speak the same public language. They may have their own, non-public, unshared sets of reasons and arguments, and these non-public argumentative "languages" may be vitally important to them. Citizens cannot be expected to shrug them off lightly, assuming that they are able to do so at all. They will in all likelihood want to bring their non-public bases of judgement to bear on important political issues. This sets the fact of pluralism and the principles of deliberative democracy at odds. For legitimation to take place, reasons must be shared; but it is a fact that citizens are moved by unshared reasons. Theorists of deliberative democracy are faced with two unpalatable solutions.

The first is to adopt what Roberto Frega calls a "gatekeeper" conception of public reason,[15] and add to the ethos of good citizenship a norm requiring citizens to refrain from drawing on their non-public beliefs in public deliberation. The second is to argue, as Paul Weithman does, that citizens may both vote and deliberate using non-public reasons.[16] Both solutions cause problems. Exclusion makes a very heavy demand on adherents of non-public doctrines, threatening to ostracise them from the political process if they do not censor their beliefs. This hardly seems consistent with the spirit of democracy. Inclusion undermines the process of legitimation through discourse in a common language, which puts into question the value of deliberation. If citizens simply swap mutually unintelligible reasons, how can a

---

[15] Roberto Frega, "Equal Accessibility to All: Habermas, Pragmatism, and the Place of Religious Beliefs in a Post-Secular Society," *Constellations* 19:2 (June 2012). Frega criticises Rawls, Habermas and Lafont for using gatekeeper conceptions of public reason, and proposes his pragmatist model of public deliberation as enquiry, where there is no fixed content to public reason, as an alternative.

[16] Weithman op. cit., especially chapter 4. Weithman does place some restrictions on the use of non-public reasons. Citizens must sincerely believe that the government would be justified in adopting the policies they vote for and advocate, and in the case of public advocacy, they must indicate why they think these policies are justified.

legitimising discourse take place?[17] Most theories of deliberative democracy which take account

of reasonable pluralism attempt to strike a balance between inclusion and exclusion. This is the

case with Rawls' late theory of political liberalism, which Habermas takes as the starting point

for his own postsecular model.

---

[17] I have stated the problem in the baldest terms. The difficulties pluralism causes for the theory of deliberative democracy are eased if we do not conceive of non-public beliefs as completely incomprehensible to non-adherents, or if we see the requirement to abstain from using them in public as a minor rather than a major burden for adherents.

3.4 Rawls' Model – Political Liberalism

In his early work, Rawls assumed that one of the characteristics of a "well-ordered society" was that "everyone has a similar sense of justice and in this respect a well-ordered society is homogenous."[18] In other words, he assumed that all members of society would be able to endorse his account of *justice as fairness*. Rawls later conceded that this was not realistic – modern societies are characterized by the existence of a multitude of *comprehensive doctrines*, meaning the religions, philosophical theories, and ethical positions in which citizens believe. (PL xviii) His mistake in *A Theory of Justice* was to present justice as fairness as itself a comprehensive doctrine which all citizens could come to embrace. (IPPR 179) In fact, "the diversity of reasonable comprehensive religious, philosophical, and moral doctrines found in modern democratic societies is not a mere historical condition that may soon pass away; it is a permanent feature of the public culture of democracy." (PL 36) Some comprehensive doctrines are reasonable, others unreasonable. (PL 63-4) Rawls calls the existence of multiple reasonable comprehensive doctrines "the fact of *reasonable pluralism*," and considers it the inevitable result of "free practical reason within the framework of free institutions." (PL 37)[19] Political theory has to take account of this fact – theorists cannot simply assume, as Rawls did in *A Theory of Justice*, that all citizens accept the same set of values for the same reasons. Rawls insists that the fact of reasonable pluralism – as opposed to pluralism tout court (PL 144)[20] – is not something regrettable. "To see reasonable pluralism as a disaster," he asserts, "is to see the exercise of reason under the conditions of freedom itself as a disaster." (PL xxvi-xxvii)

---

[18] Quoted by Joshua Cohen, "Moral Pluralism and Political Consensus," in David Copp, Jean Hampton, and John Roemer eds., *The Idea of Democracy* (Cambridge: Cambridge University Press, 1993), p 270. Cohen attributes it to the first edition of *A Theory of Justice,* p 263. The passage seems to have been excised from the 1997 revised edition, which is itself indicative of the change in Rawls' position.

[19] See PL 56-7 for a more detailed account of the "burdens of judgement" which generate reasonable pluralism.

[20] Rawls credits Cohen, op. cit., for drawing his attention to this distinction. See PR 36.

Rawls' aim in his later writings, then, was to adapt the account of democracy given in *A Theory of Justice* so as to accommodate reasonable pluralism. This means finding a place in the model for reasonable comprehensive doctrines, while excluding unreasonable ones. Rawls considers comprehensive doctrines which reject legitimate law and constitutional democracy, with its accompanying set of rights and liberties, to be unreasonable. (IPRR 132) Fundamentalist religions and authoritarian political ideologies, such as the theory of the divine right of kings, are examples. (IPRR 173) Unreasonable doctrines also reject the ideal of public reason, which I will discuss below.[21] Rawls might also consider any comprehensive doctrines unreasonable, if they are propagated or enforced in an illiberal way.[22] Any society, he says, may contain "unreasonable and irrational, even mad, comprehensive doctrines. In their case the problem is to contain them so that they do not undermine the unity and justice of society." (PL xviii-xix)[23]

The governing idea of Rawls' new model is the *duty of civility.* This is the moral obligation which citizens – particularly legislators – owe to each other to explain their political actions, their decisions and support for policies, within the terms of *public reason*. (PL 217) Public reason itself is the locus of reasons which citizens mutually understand, and can reasonably expect each other to accept. In public political discourse, citizens make use of a "family" of *political conceptions of justice* – that is, accounts of justice focused on constitutional essentials and the

---

[21] Rawls' most detailed attempt to distinguish between reasonable and unreasonable comprehensive doctrines turns on the *criterion of reciprocity*. Citizens are reasonable, and meet the criterion of reciprocity, when they are willing to offer each other fair terms of cooperation, according to what they see as the most reasonable political conception of justice, and to abide by those terms mutually. A comprehensive doctrine which embeds a political conception which allows this is reasonable, one which does not is unreasonable. See IPRR 136-8, 172-3. In essence, Rawls thinks that beliefs capable of accepting that in politics, it is what is *reasonable* that matters, not what is absolutely *true*, themselves count as reasonable. Beliefs which cannot put aside their vision of absolute truth in the political arena are unreasonable.

[22] See PL 37 n 39, where Rawls says that otherwise reasonable comprehensive doctrines, such as Kantian or Millian liberalism, could only overcome reasonable pluralism if they were enforced by oppressive state power. In such hypothetical cases, these doctrines would presumably count as unreasonable.

[23] Rawls even compares unreasonable comprehensive doctrines to "war and disease," which must be contained so that they "do not overturn political justice." See PL 64 n 19. There is a parallel here with Habermas' distinction between reflexive and unreflexive beliefs, which I discuss in §3.7.

basic structure of society, rather than comprehensive doctrines detailing the good for human beings and other metaphysical issues. (PL 213-6)[24] Rawls says that these political conception of justice should be *freestanding*. This means that they are not merely comprehensive doctrines applied to the basic structure, and are capable of being expounded without reference to any comprehensive doctrines (PL 12); nonetheless, a citizen can "embed" a political conception in her comprehensive doctrine (PL 12, 144-5), and a successful political conception will be capable of being embedded in an *overlapping consensus* of reasonable comprehensive doctrines. (PL 15)[25] Rawls describes a political conception of justice as a "module" which can "slot into" many different comprehensive doctrines. (PL 12-3)[26] The democratic policies and institutions which political conceptions justify are grounded, at a remove, in the overlapping consensus of the reasonable comprehensive doctrines in which citizens believe; this gives these policies and institutions, says Rawls, a more secure and sustainable foundation than a mere *modus vivendi*, a temporary truce between hostile comprehensive doctrines. (PL 146-7)

Rawls thinks that this arrangement allows citizens to refrain from bringing their incompatible comprehensive doctrines, with all their metaphysical breadth and subjective depth as absolute truths, into the political arena. Justice as fairness, understood now as a political conception of justice rather than a comprehensive doctrine, is naturally the political conception which Rawls favours most, but it only one amongst many. (IPRR 140-1)[27] There is also room

---

[24] See IPRR 143 for a more detailed definition.

[25] See also John Rawls, "Reply to Habermas," at PL 389. Rawls thinks that unreasonable comprehensive doctrines cannot be part of an overlapping consensus which sincerely supports constitutional democracy and public reason – they can only co-exist with other doctrines and the democratic state as part of a *modus vivendi*. See IPRR 178-9.

[26] Rawls goes into greater detail about how a comprehensive doctrine connects with a political conception of justice at PL 242 n 31.

[27] An anonymous reader has suggested to me that Rawls only envisages there being *one* political conception of justice in the society described in his later work, namely a political-not-metaphysical version of justice as fairness, which all citizens "slot into" their comprehensive doctrines. In other words, there is a reasonable plurality of comprehensive doctrines, but only one political conception of justice. I cannot see any basis for this interpretation, since Rawls says very clearly in "The Idea of Public Reason

within public reason for Habermasian views of discursive legitimacy, Catholic ideas of the common good, and so on, so long as these political conceptions of justice meet Rawls' criteria for being reasonable. (IPRR 142)

The duty of civility and the requirements of public reason apply strictly to the *public political forum*, which consists of the discourse of judges in criminal and constitutional courts, of government officials and legislators, and of candidates for office and political parties more broadly. (IPRR 133-4) It does not apply to the *background culture*, consisting of the rest of society, in which non-public reasons may have free play. (IPRR 134)[28] Of course, as Rawls notes, this is a moral rather than a statutory requirement: fulfilling the duty of civility at the correct time is part of the ethos of being a good citizen, not something which can be enforced by law. (PL 127) The duty of civility applies not only to elected officials and public figures, but also to citizens when they debate or vote about issues concerning constitutional essentials or the basic structure of society. (PL 215-6) Thus they may not make use of reasons and arguments drawn from their comprehensive doctrines when engaging in these kinds of debates. (PL 224-5) For less momentous political matters the demands of public reason are relaxed, and in the background culture they do not apply at all. (PL 245-6, 220)

Political liberalism, then, assumes a clear division between public (political conceptions) and non-public (comprehensive doctrines) sets of reasons and arguments. It would be easy to conclude that Rawls endorses a version of deliberative democracy which largely excludes non-public reasons, and by extension a strict version of secularism, since religions are comprehensive doctrines. This is not how Rawls sees matters. He is adamant that public reason, as he defines

---

Revisited" that "the content of public reason is given by a family of political conceptions of justice, *and not by a single one*… Of these, justice as fairness, whatever its merits, *is but one*." (IPRR 140-1, emphasis added.) Perhaps my reading of Rawls is more agonal than the standard interpretation, but it is evidently supported by his final articulation of his mature political philosophy.

[28] This mirrors Habermas' distinction between a secular formal public sphere and a postsecular informal public sphere. At the same time these aspects of the models are not identical, since not all comprehensive doctrines are religions.

it, is not synonymous with secular reason. The duty of civility and requirements of public reason rule out the use of *any* reasons or arguments drawn from comprehensive doctrines in the public political forum, *including non-religious comprehensive doctrines*. (IPRR 143, 147-8) It is true that overtly religious reasons and arguments are excluded from public reason; but then, so are many non-religious reasons and arguments. All religions are comprehensive doctrines, but not all comprehensive doctrines are religions. Non-religious reasons and arguments are not automatically admitted to the public political forum, simply by virtue of being secular. This is a point which Habermas overlooks in his criticisms of Rawls.

It would not in fact be accurate to say that Rawls excludes comprehensive doctrines from his model of deliberative democracy. Firstly, he lists three ways of publically using comprehensive doctrines, including religions, which do not aim directly at agreement within the terms of public reason. Thus they do not have to fulfil the duty of civility. *Declaration* means the public assertion of one's comprehensive doctrine, which may aim at demonstrating to other citizens that the doctrine can embed a reasonable political conception of justice, and can form part of an overlapping consensus. (IPRR 155) *Conjecture* means arguing from (one's understanding of) another citizen's comprehensive doctrine, in order to show them that their doctrine can embed a reasonable political conception of justice. (IPRR 155-6). *Witnessing* means expressing one's deep objection to a political decision which one regards as procedurally legitimate, but wrong or immoral in terms of one's own comprehensive doctrine. (IPRR 156 n 57) This is a case of objecting purely in terms of a particular comprehensive doctrine, not invoking political conceptions of justice in order to argue for overturning an illegitimate decision, or engaging in civil disobedience. Rawls cites Quakers peacefully expressing their opposition to military action, and Catholics to abortion, as examples of witnessing. Of these three methods of using comprehensive doctrines, declaration and conjecture have some political purpose, in that they may ultimately strengthen the ideal of public reason and belief in an overlapping consensus. None of the methods, however, works within the terms of public reason in the sense

of offering reasons and arguments drawn from political conceptions of justice, which other citizens may be reasonably expected to share.

Secondly, Rawls considers instances of the use of comprehensive doctrines which *do* directly aim at reaching agreement within the bounds of public reason. He offers two interpretations of how the requirements of public reason can be met, the *exclusive view* and the *inclusive view*. On the exclusive view, reasons drawn from comprehensive doctrines, religious and non-religious, are permanently excluded from public reason. On the inclusive view, they may be introduced on the condition that their introduction ultimately strengthens the ideal of public reason. (PL 247) As far as Rawls is concerned, the question of whether the exclusive or inclusive view is more appropriate for a particular society depends on the circumstances of that society. In a well-ordered society with a recognized and entrenched overlapping consensus, Rawls sees no reason for introducing comprehensive doctrines into public political discourse. (PL 248) Doing so would serve no purpose, since citizens know that an overlapping consensus of their comprehensive doctrines supports the various political conceptions of justice, and thus have no difficulty in honouring the ideal of public reason. Here the exclusive view should be adopted, since including comprehensive doctrines in public political discourse would do nothing to strengthen the ideal of public reason: in a very well-ordered society, public reason is already strong enough. In a less well-ordered society, where citizens have doubts about whether there is an overlapping consensus, Rawls thinks that it is reasonable to relax the duty of civility and adopt the inclusive view. The idea is that in this case, publically presenting the comprehensive doctrines which support their political conceptions of justice would reinforce citizens' belief that there is indeed an overlapping consensus. (PL 248-9) This may be achieved through "declaration," as described above. Ultimately, then, the purpose of adopting the inclusive view in the short term is to strengthen the ideal of public reason in the long term.

Rawls considers another case, that of a society which is not at all well-ordered, meaning that there are serious divisions about constitutional essentials. He uses the examples of the United States before the abolition of slavery, and during the era of racial segregation. Rawls considers a very inclusive view of public reason to be acceptable in these circumstances, provided once again that making political use of comprehensive doctrines pushes the society towards a well-ordered condition, in which the ideal of public reason can be fully honoured. (PL 250-1) Contingent conditions make this necessary: "it may happen that for a well-ordered society to come about in which public discussion consists mainly in the appeal to political values, prior historical conditions may require that comprehensive reasons be invoked to strengthen those values." (PL 251 n 41) As examples of these conditions, Rawls suggests a lack of reasonable pluralism (the society only contains a small number of strongly held comprehensive doctrines) and a lack of acceptance of the ideal of public reason and the duty of civility. It was thus reasonable, says Rawls, for the nineteenth-century abolitionists and the twentieth-century Civil Rights movement to make religious arguments against slavery and segregation. In a sense, these people were honouring the ideal of public reason by creating the conditions for it to ultimately flourish. Habermas uses similar historical examples to argue for the positive political effects of religious contributions to the public sphere (T 305); unlike Rawls, he does not place them in their historical context.

Thirdly, in "The Idea of Public Reason Revisited," Rawls modifies the rules governing the duty of civility, even within the *public political forum* of a well-ordered society.[29] He calls this the "wide view of public political culture." (IPRR 152) There is now a place for statements based on comprehensive doctrines, even within the precincts of public reason. Rawls says that

> reasonable comprehensive doctrines, religious or nonreligious, may be introduced
> in public political discussion at any time, provided that in due course proper political
> reasons – and not reasons given solely by comprehensive doctrines – are presented

---

[29] By "public political forum" Rawls means public discussions about constitutional essentials and the basic structure of society in which the ideal of public reason applies.

> that are sufficient to support whatever the comprehensive doctrines introduced are said to support. (IPRR 152, see also 144)

This is Rawls' famous proviso. It is, as he concedes, highly ambiguous. The phrase "in due course" covers a multitude of sins. It is not clear, for example, how quickly the presentation of public reasons needs to take place, or whether the person who carries it out need be the same person who made the original, comprehensive-doctrine-based argument. (IPRR 153) Habermas focuses on this second point and makes it a cornerstone of his postsecular model. Rawls simply says that the details of the proviso need to be worked out in practice.

Even with the proviso and the wide view of public political culture, Rawls sees his model as excluding "raw" comprehensive doctrines and honouring the ideal of public reason. "What we cannot do in public reason," he says, "is to proceed directly from our comprehensive doctrine, or a part thereof, to one or several political principles and values, and the particular institutions they support." (IPRR 145-6) A citizen might make a political argument by introducing the story of the Good Samaritan into public political culture, but they would then be obliged to justify their proposal "in terms of proper political values." (IPRR 146)[30]

At the same time, Rawls heavily qualifies his model's insistence on the ideal of public reason, to the extent that he introduces several means for using comprehensive doctrines, including religions, in the public political forum, and discusses several situations in which the political use of comprehensive doctrines is permissible. The duty of civility and the ideal of public reason do not apply at all to the background culture. Even within the public political forum, they are ideals, part of the ethos of good citizenship, not legal requirements. Critics of Rawls sometimes discuss political liberalism as if it was a theory of strict secularism, or even an endorsement of *laïcité.*[31] Nothing could be further from the truth. As I have shown in this

---

[30] This passage at least suggests that the same person who introduces the comprehensive-doctrine-based argument should be the one to find a proper political substitute for it.

[31] A term associated with the strong secularism of the French Third Republic's 1905 law on the separation of church and state. It has come to be associated with the principle that religion should be entirely absent from public life. While they may not use this term, Weithman, Wolterstorff and other

section, Rawls goes out of his way to give religious comprehensive doctrines a place in his model.

Nonetheless, Habermas begins the construction of his postsecular model from these criticisms

of Rawls, as we will see in the following sections.

3.5 The Split Identity Argument

Let us return to Habermas' moral arguments for postsecularism. He identifies a significant problem with secular models of deliberative democracy, such as Political Liberalism. They impose a heavy cognitive burden which is asymmetrically distributed among citizens – only religious believers suffer from it. (BNR 136) What burdens religious citizens is the obligation to contribute to the public sphere in secular language, finding their own secular "translations" (as Habermas puts it) for the religious reasons and arguments which subjectively move them. Non-religious citizens are spared this, according to Habermas – the public sphere already speaks their language. So, the principles of secular deliberative democracy force religious citizens to go through a cognitively-demanding process of "translation," which their non-religious counterparts are spared. Habermas presents this argument as an immanent critique, saying that the "liberal state contradicts itself if it demands that all citizens conform to a political ethos that imposes unequal cognitive burdens on them." (BNR 136)[32] If political equality is one of the central values of liberalism, goes the argument, then a state describing itself as liberal cannot ask more of one group of citizens than of another. In this case, it cannot demand that religious citizens, alone, undertake a complex translation process if they wish to contribute to public political discourse. Strict secularism is at odds with the liberal value of equality.

This cognitive burden leads to two unacceptable outcomes: religious citizens are obliged to *split their identities,* and feel that secular laws and policies are *not legitimate*. I will discuss the first point in this section. Requiring religious citizens to translate their own public contributions into secular language, says Habermas, is the equivalent of forcing them to "split their identities" into public (secular) and private (religious) segments. (BNR 125-6, 127) Here he is directly following Weithman, Wolterstorff and other critics of Rawls. Once again, Habermas presents his argument against strict secularism as an immanent critique of liberalism, saying that

---

[32] Habermas tends to use the adjectives "deliberative" and "liberal" interchangeably when discussing models of democracy. This is questionable, but it would be beyond the scope of this text to discuss it.

"a state cannot encumber its citizens, to whom it guarantees freedom of religion, with duties that are incompatible with pursuing a devout life – it cannot expect something impossible of them." (BNR 126) A requirement which results in identity-splitting is unacceptable because it is in conflict with one or more liberal principles – either freedom of religion, the right to personal integrity, or the right to pursue one's vision of the good life, which for religious citizens means living a devout life. A liberal political system falls into contradiction if it imposes such a requirement. Melissa Yates sums up the split identity objection as the argument that "the ideal of public reason places an unfair obligation on citizens to split themselves into non-public and public selves: the nine-to-five self speaks in public terms and keeps his or her religious commitments to him or herself, and only after hours can she or he fully express her or his true self."[33]

What Habermas means by "impossible" is ambivalent. By and large, he follows Wolterstorff's interpretation, and understands the impossibility in question as moral rather than psychological. Wolterstorff argues that for many religious believers, making public use of their beliefs is an integral part of a devout life. Religious belief is not on optional extra, a garment which can be shrugged off at the entrance to the public sphere, but an orientation which governs every aspect of life, including political action. (BNR 127-8)[34] Asking a believer to split her identity into public and private, secular and religious, aspects is to ask her to give up her identity and conception of the good. Rawls and his ilk ask too much, and go beyond the bounds of what a consistent liberal political theory may demand. On this interpretation, religious citizens are not necessarily *incapable* of understanding or using secular reasons and arguments. Habermas says that it is important to distinguish this moral objection "from the empirical observation that many

---

[33] Melissa Yates, "Rawls and Habermas on Religion in the Public Sphere," *Philosophy and Social Criticism* 33:7 (2007), p 883.

[34]Habermas is quoting Wolterstorff from Audi and Wolterstorff op. cit, p 105.

citizens who take a stance on political issues from a religious viewpoint do not have enough knowledge or imagination to find correspondingly secular justifications that are independent of their authentic beliefs." (BNR 127) On the other hand, Habermas sometimes conceives of the "impossibility" of asking religious citizens to split their identities in just this literal way. Here he follows Weithman's assertion that many religious citizens are unable to discern "any "pull" from any secular reasons,"[35] an argument which he says he finds "compelling." (BNR 128) This is an empirical claim rather than an immanent critique of liberalism – Weithman, followed by Habermas, claims that as a matter of fact, many religious citizens can make no sense of non-religious reasons and arguments. Since ought implies can, there can be no valid political norm requiring them to do so; thus strict secularism must be rejected.[36]

Habermas assumes that religious citizens alone bear a heavy cognitive burden, and suffer from the split-identity problem. "To date," he says, "only citizens committed to religious beliefs are required to split their identities, as it were, into their public and private elements." (FHN 109) As I have noted, there are non-religious comprehensive doctrines,[37] although Habermas pays little heed to them.[38] On a Rawlsian understanding of public reason, *all*

---

[35] Quoting Weithman op. cit., p 157.

[36] Catherine Audard questions this argument. She suggests that Habermas is attempting to refute an "ought" – citizens ought to limit their public deliberation to the terms of public reason – with a "can" – many of them find it "impossible" to do so. The problem with this is that the "can" is so weak. "But in what sense would this be an argument against a duty of civility, a duty to use public reasons? Habermas says that an 'ought' implies a 'can,' but would he maintain that if the 'can' is merely burdensome, it condemns the 'ought'?" See Catherine Audard, "Rawls and Habermas on the Place of Religion in the Public Domain," in James Gordon Finlayson and Fabian Freyenhagen eds., *Habermas and Rawls: Disputing the Political* (New York and London: Routledge, 2011), p 238.

[37] See for instance James W. Boettcher, "Habermas, Religion and the Ethics of Citizenship," *Philosophy and Social Criticism* 35:1 (2009), pp 233-4: "Although the difficulties associated with making a public/non-public distinction in political judgement and with setting aside the whole truth (as one sees it) are typically discussed as burdens only for religious citizens, they might be encountered by philosophical naturalists and other, non-religious citizens as well."

[38] This is reflected in his terminology. Habermas refers throughout his recent work to "secular citizens," contrasting them to "religious citizens." This is an unfortunate choice of words. We are all secular citizens, if we are citizens of secular states; and the state remains explicitly secular even in Habermas' model. It also creates the impression that non-believers alone are in favour of secularism, and that all religious citizens are opposed to it. This is far from the truth. Habermas, who often calls himself

comprehensive doctrines, religious and non-religious, are excluded from public political discourse. Their adherents must limit themselves to using political conceptions of justice, or find acceptable public substitutes for their comprehensive statements. In other words, *all* citizens suffer from identity-splitting and the cognitive burden of public reason. A sincere Marxist or Kantian, for example, would have to make a difficult mental division between public and private registers. He would have to embed a political conception of justice in his comprehensive doctrine, and go through the process of substituting acceptable public reasons for private ones. If this is the case, then we cannot speak of an *asymmetrical* cognitive burden, weighing only on the shoulders of religious citizens. On this interpretation, Habermas' first immanent critique of secular liberal democracy does not succeed – a liberal political system which imposes symmetrical burdens on its citizens is entirely consistent with its own principles. Habermas could try to salvage the identity-splitting argument by arguing that while the phenomenon affects all citizens, it is more harmful to believers than to non-believers. He would have to make the case that religions are different to all other comprehensive doctrines – more powerful, more demanding, more central to their adherents' lives than secular beliefs. There is no sign of such a dubious argument in his recent writings.

Habermas does not doubt that identity-splitting is a bad thing, but this is by no means obvious. Some strands of political philosophy acknowledge the existence of the phenomenon without characterising it as harmful to citizens. Hannah Arendt, for example, champions the distinction between private and public identities, arguing that our public political actions alone raise us out of the realm of necessity into the realm of freedom. She has no patience with the idea that only an indivisible identity is authentic.[39] Melissa Yates notes that Rawls does not

---

"religiously tone deaf," is the one arguing against strict secularism, and there is no reason why a devout believer should not also be a firm secularist. Above all, it implies that there are two homogenous, diametrically opposed blocks of citizens, the religious and the secular. The facts that non-religious citizens have sincere beliefs which may not be welcome in the public sphere, and that many religious citizens may have no trouble coping with secularism, are obscured.

[39] Hannah Arendt, *The Human Condition* (Chicago and London: University of Chicago Press, 1998). Arendt divides human actions into three categories: *labour*, carried out by *animal laborans*, which is

attempt to refute the accusation that Political Liberalism causes identity-splitting. He simply deflates it. He concedes that there is a split within a citizen's "moral identity" between the element comprising a comprehensive doctrine and the element comprising a political conception of justice. However, since these are *both* part of the moral identity, "a reasonable citizen acting according to the duty of civility should not, according to Rawls, be said to forgo his or her moral identity full stop, though he or she is expected to constrain a non-political moral identity in the public sphere."[40] This argument amounts to admitting that identity-splitting takes place, but claiming that since both the public and private "halves" of a citizen form part of their overall moral identity, it is not a matter of harmful inauthenticity or a violation of personal integrity.

The assumptions behind the split-identity objection are uncongenial to Habermas' other philosophical commitments. Rawls' critics, including Habermas, assume that a religious citizen has a unitary pre-political identity as a believer. The imposition of a secular political identity on them causes harm, since it disrupts this authentic identity – this is what makes identity-splitting a bad thing. Now, the idea that individuals are fully-formed before entering political society, and that the latter cannot legitimately modify them, is associated with social contract theorists such as Hobbes, Locke and others in the right-liberal tradition. Habermas ought not to be joining this company if he is to be consistent with his earlier work. He is more often associated with a constellation of thinkers such as Hegel, Marx, the early Frankfurt School, and many twentieth

---

brutish, repetitive and done for the sake of survival; *work*, carried out by *homo faber*, which is creative but instrumental; and *action*, carried out by the members of political communities, which is free, spontaneous and public. Politics exemplifies action. So, far from being a harmful burden, the division of our identities into public and private elements, which politics requires, is vital if we are to attain truly human freedom. See Eli Zaretsky, "Hannah Arendt and the Meaning of the Public/Private Distinction," in Craig Calhoun and John McGowan (eds.) *Hannah Arendt and the Meaning of Politics* (London and Minneapolis: University of Minnesota Press, 1997), p 223.

[40] Yates op. cit., pp 883-4. Yates in fact rejects this argument and sides with Rawls' critics. See PL 30-2 for Rawls' account of moral or non-institutional identity.

century social theorists. All of these thinkers emphasise the social nature of humanity, and in different ways criticise the right-liberal conception of a reified, pre-political human nature.

For example, Erving Goffman's concept of *dramaturgical analysis* has had a significant influence on Habermas's work. (TCAI 90-1, 93-4) Goffman argues that a person's selfhood is "presented" or "acted out" in an intersubjective process generated between performer and spectators. The exact presentation of self will vary depending on the social situation in which it takes place.[41] Habermas accepts much of this, assimilating it into his theory of formal pragmatics, while criticising Goffman's theory for being overly instrumental. This understanding of the self as a variable, intersubjective performance calls into question the idea of a unitary, pre-political self on which the split-identity argument relies. It is puzzling, to say the least, to see Habermas endorsing the latter position.

The most serious problem with the identity-splitting argument is that Habermas uses it in an inconsistent and ad hoc manner. To be consistent, he would have to hold that any norm which forces a citizen to suppress elements of her identity when engaging in public political action is unacceptable; the secular or religious nature of that identity would be irrelevant. It is clear that Habermas does not use the identity-splitting argument in this way. At one point he argues against Audi's "principle of secular motivation,"[42] and says that the link between a citizen's private motivation and his public arguments for supporting a law or policy "has no import for assessing his contribution to maintaining a liberal political culture. For in the final analysis, only the manifest reasons have institutional implications for the formation of majorities and decision-making within the relevant political bodies." (BNR 126) Habermas wants to counter

---

[41] See Erving Goffman, *The Presentation of Self in Everyday Life* (London: Penguin, 1990).

[42] Which states that a "one should not advocate or promote any legal or public policy restrictions on human conduct unless one not only has and is willing to offer, but is also *motivated by*, adequate secular reason, where this reason (or set of reasons) is motivationally sufficient for the conduct in question." See Audi (1989), p 284.

the policing of motivations which Audi advocates, but the upshot of his argument is that a disjuncture between private motivations and public reasons and arguments *is acceptable*. In a recent text, Habermas replies to Lafont's argument that his postsecular model forces non-religious citizens to act insincerely towards their religious counterparts,[43] by saying that this objection "has no force for a discourse-theoretical understanding of the democratic process… only public utterances, hence actual contributions to the formation of opinions and consensus-building, *and not mindsets*, have a bearing on the legitimizing power of democratic discourses."[44] In the case of non-religious citizens, then, only their actual public contributions matter; the mismatch between their public utterances and their subjective mindsets – surely a case of identity-splitting – is irrelevant to deliberative democracy. So, is identity-splitting acceptable or not? If Habermas can dismiss Lafont's argument that non-religious citizens suffer by having their identities split, where does that leave his own argument about religious citizens?

Identity-splitting affects non-religious as well as religious citizens, and as such cannot be described as an asymmetrical cognitive burden. It may be an unproblematic (Rawls) or even a desirable (Arendt) phenomenon. The argument that it is harmful seems to rely on a naïve account of personal identity which Habermas would otherwise reject. Finally, Habermas uses it in an inconsistent and ad hoc way. The split identity argument seems like a weak prop for the postsecular model. Let us now consider Habermas' other argument for the model, and see if it is any stronger.

---

[43] In Lafont op. cit., a revised version of which can be found in *Habermas and Religion*, pp 230-248.

[44] Jürgen Habermas, "Reply to My Critics," in *Habermas and Religion* p 372 n 40, emphasis added.

3.6 The Legitimation Crisis Argument

Habermas borrows the split identity argument from critics of Rawls such as Wolterstorff and Weithman. The legitimation crisis argument is his own, and is tied closely to the account of legitimate lawmaking he outlines in *Between Facts and Norms.* The discourse theory of law and democracy gives Habermas grounds to argue that the exclusion of religious contributions from political discourse robs the resulting laws and policies of legitimacy, in the eyes of religious believers at least. A strictly secular deliberative democracy, then, will appear illegitimate to many of its citizens – it will experience a legitimation crisis, which at its worst can lead to violence. (FHN 102-3)

This argument takes as its starting point Habermas' understanding of legal legitimacy. Legitimacy per se is summed up in the *discourse principle*, (D):

> Only those norms can claim to be valid that meet (or could meet) with the approval of all affected in their capacity *as participants in a practical discourse*. (MCCA 66)

This applies to all action norms. Legal legitimacy is a subset, or a particular instantiation, of legitimacy as described by the (D) principle. A legitimate law is not identical to a legitimate moral norm, for Habermas, but must be generated and enacted within a certain institutional framework. Legal legitimacy is summed up in the *principle of democracy*:

> only those statutes may claim legitimacy that can meet with the assent of all citizens in a discursive process of legitimation that in turn has been legally constituted. (BFN 110)

The principle of democracy is generated by applying (D) to the *form of law*.[45] By "the form of law," Habermas means the constitution of legal persons as rights-bearers. Moral norms apply to human beings per se, but laws apply to human beings in their capacities as legally defined rights-bearers. Habermas divides the rights in question into five categories, collectively known as the

---

[45] See James Gordon Finlayson and Fabian Freyenhagen, "Introduction" to *Habermas and Rawls: Disputing the Political*, pp 8-12, for an overview of these points.

system of rights, which define legal personhood.[46] This system of rights ensures that legal persons are autonomous, have equal measures of basic liberties, equal guarantees of security, and, crucially, equal access to the public sphere. (BFN 122-3) It is the "interpenetration" of the system of rights and the form of law, summed up in the democracy principle, which sets the conditions for legitimate lawmaking, according to Habermas.

The essence of this theory is that when citizens' contributions to the fora of the *informal public sphere* (made under the correct discursive conditions, within the correct constitutional framework) pass as input into the *formal public sphere* of the state and the legal system, legitimate laws will be produced as output. (BFN 371-2, 441-2) The informal public sphere is an unregulated space, consisting of the media, civil society and public discussions in general; the formal public sphere is a regulated space, containing the state, parliament and legal system (Habermas sometimes calls it the "parliamentary complex"), from which laws and policies emerge. (BFN 307-8) Habermas describes the informal public sphere as "besiege(ing)" the "fortress" of the formal, (BFN 486-7) and of the two being connected by a series of "sluices." (BFN 354-8) Legal legitimacy depends on both spheres interacting correctly. The formal public sphere ensures that the system of rights is in place, allowing citizens to make contributions to public discourse in an unforced, undistorted way. The informal public sphere feeds spontaneous popular impulses into the parliamentary complex, breathing life into its products. Finally, the formal public sphere tames the "wild" impulses of the informal, and transmutes them by imposing the legal form upon them.

One of Habermas' formulae for describing legal legitimacy is his assertion that citizens can understand themselves as *both the authors and addressees* of legitimate laws. (BFN 120,

---

[46] Hugh Baxter, *Habermas: The Discourse Theory of Law and Democracy (Jurists: Profiles in Legal Theory)* (Stanford: Stanford University Press, 2011) offers a more in-depth account of the theory. On this point, see pp 70-5.

passim)[47] If a citizen can see that the reasons and arguments they presented in the public discussion about a particular issue have fed into a law about that issue, they can see themselves as both the author and the addressee of that law, and thus accept it as legitimate and comply with it. Habermas must be using a weak sense of "authorship" here. Since universal assent is very rare, outvoted minorities and those who lose arguments in the informal public sphere must still be able to see themselves as the "authors" of laws which they opposed. Presumably it is enough to acknowledge that the dispute to which you contributed went on to influence the process of lawmaking, even though you came out on the losing side. Alternatively, Hugh Baxter suggests that endorsement of the procedures of argumentation and lawmaking on the part of outvoted minorities replaces endorsement of the substantive outcomes of the procedures.[48]

Habermas thinks that citizens can comply with laws for two discrete reasons. It makes sense to comply *strategically* with a law if breaking it leads to sanctions. On its own, strategic compliance is not enough – it operates at the level of preconventional morality, and implies that citizens should break laws whenever it is advantageous for them to do so, if they can avoid punishment. It would also make sense to comply strategically with illegitimate laws, if they were strongly enforced. *Performative* compliance, where laws are obeyed because it is morally right to do so, is also necessary. This ensures that citizens obey legitimate laws even when they are not forced to. (BFN 448) Here we can hear an echo of Habermas' distinction between social and system integration. (TCAII 150) Citizens of modern democracies must be able to comply with laws in both a strategic and a performative manner; this is what Habermas means by saying that laws are both "facts" and "norms."

---

[47] This of course is the familair idea of autonomy or positive freedom, articulated at the individual moral level by Kant and at the collective political level by Rousseau. Habermas believes that his theory recognises the "co-originality" of positive and negative freedom, by drawing out the mutual dependence of the system of rights and popular sovereignty.

[48] Baxter op. cit., p 100.

If citizens cannot contribute to the public sphere as they see fit, if the contributions they make have little meaning for them, the process of legitimation will not take place. Citizens will see themselves as merely the addressees of the law, not its authors. This puts them in a position of heteronomy in relation to the formal public sphere, and encourages them to comply in a purely strategic manner to its laws. This is the position of religious believers in secular states, according to Habermas. (BNR 130) Strict secularism prevents them from contributing to the public sphere with the religious reasons and arguments which appeal to them. They may remain silent, or they may be forced to "translate" their contributions into secular terms. Either way, the disjuncture between their public and private modes of discourse interrupts the process of legitimation. This is the basis of Habermas' legitimation crisis argument for postsecularism. Secular deliberative democracy lacks legitimacy in the eyes of many of its citizens; it must be modified, since "those who are neither willing nor able to separate their moral convictions and vocabulary into profane and religious strands must be permitted to participate in political will formation even if they use religious language." (EFP 76)

Habermas' argument may be clearer if it is looked at from the standpoint of acknowledgement. The discourse theory of law and democracy tacitly assumes that for legitimation to take place, citizens must be able to acknowledge their discursive inputs, at both ends of the legislative process. I will call this the *acknowledgement requirement.* A citizen must be able to acknowledge her contributions to the informal public sphere as they pass through into the formal public sphere, and as they re-emerge from the latter in the guise of laws, in order to recognise those laws as her own – to see herself as both an author and an addressee. These two moments could be called "pre-legislative" and "post-legislative" acknowledgement, with the former being prior to the latter. A citizen who cannot acknowledge his contributions as they enter the formal public sphere will not be able subsequently to acknowledge the laws which

they engender.[49] Citizens who lose arguments in the informal public sphere can still experience pre-legislative acknowledgement; we must assume that this is enough to generate legitimacy. Habermas' legitimation objection to secular deliberative democracies can be cast in acknowledgement terms: strict secularism prevents religious citizens from engaging in pre-legislative acknowledgement of their discursive inputs, thus robbing the resulting laws of post-legislative acknowledgement, and legitimacy.[50] If Habermas' new postsecular model is to solve the problem of legitimation, it must enable religious citizens experience (at least) pre-legislative acknowledgement of their contributions to the public sphere.

Habermas objects to secular models of deliberative democracy using the split identity and legitimation crisis arguments. Let us now consider his solution to these problems, the model of postsecular deliberative democracy.

---

[49] In assuming that post-legislative acknowledgement can and must take place, I am assuming that the legislative process does not transform discursive inputs so drastically that the laws they become unrecognizable to their original utterers. But then, Habermas seems to assume this as well. My "acknowledgement requirement" simply fleshes out his formula of "authors and addressees."

[50] This fits in with Weithman's strong interpretation of the cognitive difficulties of religious citizens, which Habermas sometimes favours. If they are unable to feel any "pull" from secular reasons and arguments, then they will evidently fail to experience pre-legislative acknowledgement in a model which only allows secular inputs to pass into the formal public sphere. However, it can also be made to fit with Wolterstorff's interpretation, if the relevant sense of "acknowledgement" is weakened. Religious citizens may be able to understand and feel the force of secular reasons and arguments, including their own unwilling translations, but may not be able to acknowledge such inputs as *their own*. In that sense, they do not experience pre-legislative acknowledgement.

3.7 Habermas' Model – Cognitive Adjustments

To begin with, Habermas calls for a change in mentality on the part of both religious and non-religious citizens. Both groups must take part in a "complementary learning process," (BNR 111) which will result in their becoming reflexive about their beliefs. (BNR 144) These cognitive adjustments play two roles for Habermas. Firstly, they make possible the changes to deliberative procedures which are the focus of the next section, namely the introduction of an "institutional translation proviso" into the public sphere. Secondly, combined with these new procedures, they equalise the cognitive burden borne by religious and non-religious citizens.

Religious citizens, says Habermas, must adjust their beliefs to meet the three sources of cognitive dissonance which are unavoidable in a modern society:

(1) Pluralism of beliefs – believers must come to terms with the presence in society of beliefs different to theirs, including atheism and agnosticism.

(2) Scientific knowledge – believers must accept that the natural sciences have the monopoly on factual knowledge.

(3) Secular politics – believers must accept the secular nature of the constitutional state, which is neutral between worldviews and operates according to the standards of postconventional morality. (BNR 137)

Religious communities which have made these adjustments, which deserve to be called "reasonable" in Rawls' sense, (FHN 104 n 2) would naturally abstain from external violence and internal oppression. Habermas seems to regard these three adjustments as *conditions* for participating in the public sphere: reflexive religious citizens are allowed to participate, unreflexive ones are not.[51] Catherine Audard interprets them as being "conditions for entry into

---

[51] This would cohere with Habermas' earlier view of religious fundamentalisms as "incompatible with constitutional democracy." See "Struggles for Recognition in the Democratic Constitutional State," in *The Inclusion of the Other: Studies in Political Theory* (Cambridge, Massachusetts: MIT Press, 2001), p 224.

democratic societies,"[52] and this is borne out by Habermas's comment that it is in the interests

of religious communities to become reflexive, since "they thereby gain the opportunity to

exercise influence of their own on society as a whole through the political public arena." (BNR

112) Presumably they would have no such opportunity if they did not become reflexive, meaning

that reflexive and unreflexive forms of belief play approximately the same role in Habermas'

model as reasonable and unreasonable comprehensive doctrines do in Rawls'.

Habermas is not so precise about the cognitive adjustments which he expects from non-

religious citizens.[53] They are expected to become reflexive about the limits of secular reason, to

"determine the relation between faith and knowledge *self-critically* from the perspective of

secular knowledge." (BNR 309-10) The upshot of this is that non-religious citizens must treat the

disagreements between themselves and believers as *reasonable* disagreements, between

equally reasonable comprehensive doctrines, which means that "religious convictions [must be]

accorded an epistemic status that is not merely irrational from the perspective of secular

knowledge." (BNR 264) In effect, non-religious citizens must become agnostics for the purpose

of public political discourse. Habermas says that they "may neither deny that religious

worldviews are in principle capable of truth nor question the right of their devout fellow-citizens

to couch their contributions to public discussions in religious language." And crucially, a "liberal

political culture can even expect its secular citizens to take part in the efforts to translate

relevant contributions from religious language into a publically intelligible language." (BNR 113)

I will discuss the procedure of translation in the next section. Suffice to say that it is a demanding

procedure which non-religious citizens are obliged to carry out, from which religious citizens are

---

[52] Audard, "Rawls and Habermas on the Place of Religion in the Public Domain," p 244. Norbert
Brieskorn considers the three cognitive adjustments to be "conditions of entry into the public arena."
See "On the Attempt to Recall a Relationship," AWM 33.

[53] Baxter lists four cognitive adjustments which Habermas requires from non-religious citizens: (1) not to
assume, in their role as citizens, that religious reasons or arguments are automatically devoid of truth,
(2) to accept that religion has a permanent place in modern societies, and in politics, (3) to accept the
right of religious citizens to contribute to the public sphere in religious language, and (4) to take part in
translation. See Baxter op. cit., p 205.

exempt. The shared challenge of becoming reflexive, plus the unshared challenge of translation, makes the cognitive burdens endured by non-religious citizens approximately equal to those of religious citizens, according to Habermas. So the measure of this aspect of his postsecular model lies in its ability to equalize burdens.[54]

It is surprising to hear a professed fallibilist such as Habermas predicting the outcome of a learning process, and basing a political model on it. He himself expresses unease with the teleological cast of the process. He notes that "these changes in mentality count as complementary "learning processes" only from the perspective of a specific normative self-understanding of modernity," and that that self-understanding could only be grounded by invoking a (debatable) theory of social evolution. (BNR 144) Habermas is aware that in practice, the development of religious and non-religious mindsets is an open-ended process. There is no guarantee that non-believers will develop a reflexive rather than "a scientistic form of secularism," or that believers will arrive at a form of belief which is reflexive by Habermas' definition – only religious communities themselves "can decide whether a 'modernized' faith is still the 'true' faith." (BNR 145) So on the one hand, the outcome of spontaneous processes of mutual learning between religious and non-religious citizens is uncertain. On the other hand, it is beyond the power of the state to engineer such a fine-grained adjustment of attitudes in its citizens: "the requirement of complex mentalities highlights an improbable functional imperative whose fulfilment the liberal state can scarcely influence through the legal and administrative means at its disposal." (BNR 144) I take this to mean that Habermas thinks it is not possible for the state to thread the cognitive needle which stitches his model together; but

---

[54] Commentators are divided over whether it achieves this goal. Amy Allen thinks that Habermas's model still places a heavier burden on religious than on non-religious citizens. See Amy Allen, "Having One's Cake and Eating It Too: Habermas's Genealogy of Postsecular Reason," *Habermas and Religion* pp 150-1. Cristina Lafont, on the other hand, thinks that the postsecular model penalises non-believers, shifting the burden of "cognitive dishonesty" previously borne by believers onto their shoulders. See Cristina Lafont "Religion in the Public Sphere: Remarks on Habermas's Conception of Public Deliberation in Post-Secular Societies" *Constellations* 14:2 (2007), pp 248.

he also argues that it is not permissible for the state to try to do so. (EFP 75) Given these concessions, it seems that the postsecular model rests on highly contingent foundations. There is no guarantee that religious and non-religious will cultivate the necessary reflexive attitudes, and little scope for the state to encourage them to do so.

There is also an air of inconsistency about this aspect of the model. Habermas wants to solve the problems of religious believers who feel that secular processes of deliberation produce illegitimate laws, and who cannot tolerate the supposed identity-splitting required by secularism. As I noted in section 3.5, the split-identity argument seems to rest on a naïve conception of pre-political identity which sits awkwardly Habermas' other philosophical commitments. He thinks that identity-splitting is objectionable because pre-political identities, formed by religious beliefs, are integral, and must not be interfered with by the norms of the political system. And yet, Habermas' postsecular model requires citizens to make dramatic changes to their beliefs, and therefore to their identities. If the mandated (but not legally enforced) cognitive adjustments of secularism are unacceptable, why are the mandated (but not legally enforced) cognitive adjustments of postsecularism acceptable? A case could be made that the impositions of Habermas' model are in fact *more* onerous than those of political liberalism. Rawls' model requires citizens to refrain from voicing their beliefs when deliberating on matters of basic justice or constitutional essentials, but in other respects it leaves those beliefs alone. Habermas' model requires them to *change their manner of belief*.

Several commentators have drawn attention to these problems. Catherine Audard contrasts Rawls' model, which requires only a consensus of "minds" among citizens, to Habermas', which requires "hearts" as well as "minds."[55] Others have questioned Habermas' assumption that the three cognitive adjustments would be acceptable to religious believers. Amy Allen notes that Habermas' conception of reflexive belief "may make sense for theologians,

---

[55] Audard "Rawls and Habermas on the Place of Religion in the Public Domain," pp 243-4.

[but] it has little in common with the ways in which most religious believers experience their own faith." (HR 151) Habermas may not realise just how much he is asking, when he requires believers to accept the monopoly of the natural sciences over factual knowledge. This could easily be seen as a demand to downgrade sacred texts to the level of fictions – a step too far for most believers. Thomas McCarthy observes that not all religious citizens will accept Habermas' criteria of reflexive belief. Some will have a fideist perspective, which values faith and commitment above enquiry and the revisable outcomes of rational discourse. Even reflexive believers, McCarthy suggests, may hold that "modern law and secular morality can only ever be of secondary importance in comparison to the dictates of religious belief, practice, and conscience."[56]

The cognitive adjustments of Habermas' model raise as many questions as answers. What I will focus on, however, is the political model of postsecular deliberative democracy which these adjustments ground. This will be the focus on the next section.

---

[56] Thomas McCarthy, "The Burdens of Modernized Faith and Postmetaphysical Reason in Habermas's 'Unfinished Project of Enlightenment'," *Habermas and Religion*, p 118.

3.8 Habermas' Model – The Institutional Translation Proviso

Habermas modifies his discourse theory of law and democracy by distinguishing more sharply between the formal and informal public spheres, and introducing an *institutional translation proviso* (ITP) as a filter between them. In this new postsecular model, the formal public sphere remains strictly secular. Untranslated religious reasons must not be allowed to enter, since "all *enforceable* legal norms must be capable of being formulated and publically justified in a language intelligible to all of the citizens." (EFP 76)[57] So public officials, judges, and legislators are required to put aside their religious beliefs, as in Rawls' model. (BNR 128) Habermas even suggests that the speakers of parliaments should have the power to strike any explicitly religious statements from the official record. (BNR 131)

In the informal public sphere, on the other hand, religious citizens may make contributions to discourse in religious language. Non-religious citizens, as we have seen, may not object to such contributions, or automatically dismiss them: "Insofar as they act in their role as citizens, secularized citizens may neither fundamentally deny that religious worldviews may be true nor reject the right of devout fellow-citizens to couch their contributions to public discussions in religious language." (BNR 310)[58] This is more than toleration, says Habermas – it is "a self-reflexive overcoming of a rigid and exclusive secularist self-understanding of modernity." (BNR 138) He assumes that the cognitive adjustments described in the previous section make this division of labour possible. (BNR 111) Reflexive non-believers would not object to religious statements in the informal public sphere, and reflexive believers would not object to the ITP preserving the secular character of the formal public sphere. Citizens of faith, says

---

[57] As Cooke and others have pointed out, Habermas still assumes, even in his writings on the postsecular model, that "secular" and "mutually intelligible" are synonyms. See Maeve Cooke, "A Secular State for a Postsecular Society? Postmetaphysical Political Theory and the Place of Religion," *Constellations* 14:2 (2007), p 230.

[58] This stipulation is the basis of Lafont's objection. See Lafont (2007), pp 252-3.

Habermas, "may make public contributions in their own religious language only subject to the translation proviso." (BNR 132)

Religious statements are translated into secular language in the informal public sphere. The ITP blocks untranslated religious statements from passing through into the formal public sphere, and allows in only translated ones: "the institutional threshold between the 'wild' political public sphere and the formal proceedings within political bodies also functions as a filter that allows only secular contributions from the Babel of voices in the informal flows of public communication to pass through." (BNR 131)[59] Habermas often refers to the translation process as a "cooperative task" involving both religious and non-religious citizens. (BNR 131-2) This implies that both groups of citizens are obliged to work together. However, Habermas objects strongly when other theorists suggest that religious citizens have an obligation to translate (or replace, or make secular substitutions of) their own contributions to the public sphere. He criticises not only Rawls (P 25) and Robert Audi, but also Rainer Forst for imposing such an obligation. (BNR 132 n 27) His postsecular model is based on the assumption that religious citizens should *not* be required to translate their own contributions. Many of Habermas' comments make it clear that religious citizens may contribute "monolingually" to the informal public sphere, whether because they are unable to find secular translations for their contributions or because they are unwilling to do so. As such they are *not* obliged to collaborate

---

[59] Several commentators have criticised the sharp distinction between the formal and informal public spheres. Such a division seems to go against the spirit of Habermas' deliberative conception of democracy, where continuous interaction between the public sphere and the "parliamentary complex" is the key to legitimate law- and policy-making. The ITP is a porous boundary, of course, but it is a boundary nonetheless. On this point see Mariano Barbato and Friedrich Kratochwil, "Towards a Post-Secular Political Order?" *European Political Science Review* 1:3 (2009), pp 333-4. James Boettcher highlights another aspect of this problem. It is a basic feature of any model of democracy that citizens should be able to hold public officials to account. But if public officials, operating in the formal public sphere, speak a secular language which is foreign or even incomprehensible to religious citizens, operating in the informal public sphere, how can the latter hold the former to account? Moreover, in Habermas' model, one of the duties of public officials would be to ensure that the ITP is being observed – that only secular translations of religious contributions pass into the formal public sphere, and that the latter remains strictly secular. If religious citizens are allowed to operate "monolingually," how can they assess whether officials have performed this duty correctly? See Boettcher op. cit., pp 225-7.

with non-religious citizens to produce translations – their role may be limited to providing religious inputs for processes of translation carried out by non-religious citizens. Habermas says that they should "be allowed to express and justify their convictions in a religious language even when they cannot find secular 'translations' for them," (BNR 130) and that "those who are neither willing nor able to separate their moral convictions and vocabulary into profane and religious strands must be permitted to participate in political will formation even if they use religious language." (EFP 76) It follows that non-religious citizens *are* obliged to produce secular translations of religious statements. Habermas says that a "liberal political culture can… expect its secular citizens to take part in the efforts to translate relevant contributions from religious language into a publically accessible language," (BNR 112-3) and that "the 'monolingual' contributions of religious citizens then depend on the translational efforts of cooperative fellow citizens if they are not to fall on deaf ears…" (P 26) In practice, then, non-religious citizens have an obligation to produce translations, while their counterparts may contribute to the informal public sphere in exclusively religious terms.

As for the process of sacred-to-secular translation itself, Habermas says very little about it. As I have noted, he only gives one concrete example. In "Faith and Knowledge," Habermas describes an argument against stem-cell research made by Christian groups in Germany as a translation, albeit "an unfortunate one," of a religious concept into "the secular language of the constitution." (FHN 109) The statement from Genesis 1:27 that

(1) "So God created man in his own image, in the image of God created he him"

translates into the statement that

(2) "A gamete fertilized ex utreo is a subject of human rights." (FHN 114-5)

Habermas seems to consider this particular translation "unfortunate" only because these religious citizens were forced to carry it out on their own – he is talking about an unmodified

secular public sphere – not because the translation itself is unsuccessful. The fact that he returns to it several times in his recent writings suggests that he does not consider it a bad example of translation on its own terms.[60] So the process of turning (1) into (2) is an example of sacred-to-secular translation – the *only* concrete example which Habermas gives us. This is the process, taking place in the informal public sphere, which drives the postsecular model.

To summarize: Habermas' model of postsecular deliberative democracy consists of a formal public sphere which is strictly secular, and an informal public sphere in which religious reasons and arguments are permitted. In the informal public sphere, religious contributions are translated into secular language before being fed into the formal public sphere. Reflexive religious citizens may contribute to the informal public sphere in religious language alone; reflexive non-religious citizens are obliged to help them produce secular translations of their contributions. The cognitive adjustments which all citizens have acquired through mutual learning make all this possible. The ITP acts as a filter between the two public spheres, allowing translated religious statements to pass through but blocking untranslated ones.

Habermas presents his model as the solution to the problems he identifies in secular deliberative democracy: unequal cognitive burdens, identity-splitting, and de-legitimation. Postsecular deliberative democracy "release[s] religious citizens from the burden of having to make a strict separation between secular and religious reasons in the political public arena when they experience this as an attack on their personal identity," and in this way solves the problem of identity-splitting. "Religious citizens can certainly acknowledge this 'institutional translation proviso' without having to split their identity into public and private parts the moment they participate in public discourses." (BNR 130) This may be true for the most part, but it ignores the fact that some religious citizens operate in the formal public sphere, as legislators, judges or

---

[60] See for example "Prepolitical Foundations of the Constitutional State?" (BNR 110) and "A Conversation About God and the World: Interview with Eduardo Mendieta," in Jürgen Habermas, *Religion and Rationality: Essays on Reason, God and Modernity* (Cambridge: Polity, 2002).

public officials. They are still obliged to split their identities. This raises questions about Habermas' model. If public officials who are religious believers can be expected to cope with these conditions, then why not ordinary religious citizens? Public officials are, of course, held to a higher standard than ordinary citizens, but in a democracy they are assumed to be cut of the same cloth. The postsecular model, then, reduces rather than eliminates identity-splitting.[61] I have already discussed Habermas' argument that mutual learning, the process of becoming reflexive, helps to equalise the cognitive burdens borne by believers and non-believers. The differential burden of translation, which seems to fall on non-religious to a greater extent than on religious citizens, balances their burdens even more closely. (BNR 143)

The postsecular model tackles the legitimation problem by enabling religious citizens to know that their contributions to public discourse, suitably translated, are being fed into the formal public sphere. Thus they can see themselves as both the authors and addressees of the laws which emerge from it, making the laws legitimate in their eyes. They can obey them in performative as well as strategic ways. As I noted in section 3.6, Habermas' theory requires both kinds of compliance for a law to be considered legitimate. This argument assumes that secular translations of religious statements can meet the requirement of pre-legislative acknowledgement, as I have called it. Religious citizens must be able to acknowledge the translated versions of their contributions as their own, as they pass through the ITP into the formal public sphere. Habermas seems to assume just this when he says that a religious citizen needs only "the epistemic ability to consider [their] own religious convictions reflexively from the outside and to *connect them with secular views*" (BNR 130, emphasis added) in order to

---

[61] I put this question to Habermas during an e-mail exchange in September 2014. He replied that the issue only applied to MPs and candidates for political office, rather than civil servants. In the case of MPs who are believers, "they are (according to my scenario) not allowed to act as 'monolingual' religious people. They must have learned to express themselves in a secular language. But that should be no problem for politicians who must attract voters, these must be multilingual for the job anyway."(Jürgen Habermas, private e-mail, 9/9/2014.)

accept the ITP. Supposing that this is the case, monolingually religious citizens will no longer be estranged from the process of legitimation:

> Even if the religious language is the only one they speak in public, and if religiously justified opinions are the only ones they can or wish to contribute to political controversies, they nevertheless understand themselves as members of a *civitas terrena*, which empowers them to be the authors of laws to which they are subject as addressees. They may express themselves in a religious idiom only on the condition that they recognize the institutional translation proviso. Thus the citizens, confident that their fellow-citizens will cooperate in producing a translation, can understand themselves as participants in the legislative process, although only secular reasons count therein. (BNR 130-1)

The procedure of translation, then, is the key to postsecular deliberative democracy. Habermas uses it to solve the problem of de-legitimation, and as part of the solution to the problems of identity-splitting and unequal cognitive burdens. Although crucial to the model, it is a notoriously ill-defined procedure. In the following sections I will examine sacred-to-secular translation, and argue that it is the fatal flaw in Habermas' model.

3.9 Translation and Appropriation

What does Habermas mean by "translation"?

The procedure is central to his model of postsecular deliberative democracy, and yet he never gives a clear account of it. A number of theorists have expressed puzzlement about how the procedure works[62]; others have questioned the suitability of the term "translation" to describe it.[63] In this section I will review some of the interpretations of sacred-to-secular translation, identify Habermas' definition of it, and argue that his definition suffers from serious problems.

Commentators on Habermas' work have interpreted the translation procedure in various ways. Maeve Cooke equates translation with what I call ethical appropriation, and emphasises the literary re-presentation which religious stories and images require if they are to function as ethical inputs for postmetaphysical thinkers.[64] Phil Ennis sees sacred-to-secular translation as a procedure of abstracting insights from ethical discourse so that they can function in moral discourse. Ethical values, he argues, must be removed from the particular lifeworld contexts in which they are entangled, and made abstract and general, so that they can pass the universalization test of the (D) principle.[65] James Boettcher argues that the procedure of translation in Habermas' model is identical with the procedure of substitution in Rawls' model.

---

[62] See Austin Harrington, "Habermas and the 'Post-Secular Society,'" *European Journal of Social Theory* 10:4 (2007), p 552, Boettcher op. cit., pp 228-9, 230-1, Cooke (2011) pp 481-2.

[63] See Camil Ungureanu, "The Contested Relation Between Democracy and Religion – Towards a Dialogical Perspective?" *European Journal of Political Theory* 7:4 (October 2008), p 408 n 12, 417 n 46, Walhof op. cit., pp 233-4, Javier Aguirre, "Habermas' account of the role of religion in the public sphere: A response to Cristina Lafont's critiques through an illustrative political debate about same-sex marriage," *Philosophy and Social Criticism* 39:7 (September 2013), pp 643-4.

[64] Cooke (2007) p 226, and Cooke (2011), passim. She uses the terms "translation" and "salvaging" interchangeably. See § 2.5 for a more detailed account of ethical appropriation.

[65] Phil Ennis, "Habermas, Reason, and the Problem of Religion: The Role of Religion in the Public Sphere," *The Heythrop Journal* 48:6 (2007), and especially Phil Ennis, "Habermas, Democracy and Religious Reasons," *The Heythrop Journal* 51:4 (2010), pp 585-6. Ennis argues in his second article that translation should not in fact be necessary for religious inputs to play a role in the public sphere.

Rawls talks about citizens slotting political conceptions of justice into their comprehensive doctrines, and substituting reasons drawn from the former for reasons drawn from the latter ("in due course") when engaging in public political discourse; Habermas talks about citizens translating religious reasons and arguments into secular equivalents when they contribute to the formal public sphere. To all intents and purposes, says Boettcher, the two procedures are the same.[66]

Other theorists reject the vocabulary of translation entirely, and construct alternate versions of Habermas' procedure. Camil Ungureanu proposes a model of postsecular deliberative democracy based on a procedure of "selective democratic interpretation," which turns religious motives into "secular democratic justifications,"[67] rather than any variety of translation. Roberto Frega, as I have noted, argues that Habermas makes the same mistake as Rawls, in that he sets up a fixed definition of public reason as a "gatekeeper" – he uses the ITP to maintain the secular nature of the formal public sphere. Frega favours a pluralized public reason, consisting of many normative practices with their own epistemic standards, governed by a procedure of experimental enquiry along pragmatist lines.[68]

So, how does Habermas understand sacred-to-secular translation? A survey of his recent writings shows that he does not differentiate between sacred-to-secular translation and the processes of postmetaphysical appropriation. Cooke's interpretation is closest to the truth.[69]

---

[66] Boettcher op. cit., pp 228-9, 231. To be precise, Boettcher says that a weak interpretation of translation is virtually identical with Rawlsian substitution, and suggests that Habermas himself adheres to this weak interpretation.

[67] Ungureanu op. cit., p 417.

[68] Frega op. cit.

[69] It should be noted that neither Habermas nor Cooke distinguish as fastidiously as I do between conceptual and ethical appropriation. As far as they are concerned, all the varieties of appropriation, and translation, are the same. One difference between their accounts of appropriation is that Cooke equates the procedure entirely with ethical appropriation, while Habermas, as I will show, frequently equates it with conceptual appropriation. I have already argued that conceptual and ethical

When he gives historical examples of sacred-to-secular translations, Habermas typically refers to conceptual appropriations carried out by philosophers. But his chief (in fact, his only) example of the procedure appears to be an ethical appropriation. Habermas says that "a gamete fertilized ex utreo is a subject of human rights." is a secular translation of "God created man in his own image" Friedo Ricken analyses this transformation in his essay "Postmetaphysical Reason and Religion," (AWM 51-8) and a close reading of his analysis shows that it is, in my terms, an ethical appropriation of the second type[70]. The original Biblical statement, he says, contains two assertions:

(a) that human beings bear a likeness to God, and

(b) that they are creatures of God, created by God.

Philosophy expresses (a), says Ricken, in these terms:

(a') humans "are beings endowed with and bound by freedom. Because every human being is an image of God, all human beings are equally free." (AWM 52)

In the original religious version of (a), then, it is implied that

(i)   there is a God,

(ii)   this God possesses freedom, and

(iii)   humans resemble God in respect of being free.

In (a'), assumptions (i) and (ii) have dropped out. We might say that the original substance of piety enclosing the appropriated content has been stripped away. A modified version of (iii)

---

appropriations should be distinguished from each other; at the end of this section, I will argue that neither process can be synonymous with sacred-to-secular translation.

[70] In type one ethical appropriation, an ethical value plus its "substance of piety" or lifeworld context is transplanted into postmetaphysical discourse. In type two ethical appropriation, the ethical value is stripped of its substance of piety, which is replaced by a new one once the value has been taken over into postmetaphysical discourse. The first type of appropriation is suitable when the lifeworld context of the appropriators is sufficiently similar to that of the appropriated value; the second type is needed when the context of the value is remote. See § 2.5.

remains, in the form of the assertion that human beings possess freedom. We might say that (iii) is the ethical insight which is being appropriated. A new element is introduced in (a'), the assertion that

(iv)   *all* human being are *equally* free.

This is an egalitarian extrapolation of the various elements of (a). The egalitarian tone of (iv) suits a modern, secular or postmetaphysical consciousness better than the heteronymous tone of (i) and (ii). All religious assertions of this sort, however benign, are coloured by a fundamental heteronomy: humans are ontologically posterior to, or existentially dependent on, God. This is unproblematic for religious or metaphysical thinkers, but less acceptable for secular or postmetaphysical thinkers. We might say that the appropriated ethical insight is being surrounded by a new substance of piety, which will make it persuasive and effective for those who do not share the lifeworld context in which it originally appeared.

Assertion (b), says Ricken, must first be cast in negative terms: saying that human being are creatures of God means that human beings are *not* creatures of *other human beings*, they "do not owe their natural essence to other human beings." (AWM 52) Thus it becomes, in its original Biblical version, a prohibition. As Ricken puts it, quoting Habermas:

(b) anyone "who determines, '*at his own discretion*, the natural essence of another human being' [FHN 115] transgresses, in the language of the Bible, the boundary between creature and creator." (AWM 52)

Philosophical language expresses (b), says Ricken, in these terms:

(b') any human being who determines the natural essence of another "fails to respect the reciprocal relation among human beings implied by equal freedom; by determining the natural essence of the other at his own discretion, he violates the principle of equal freedom." (AWM 52)

As above, assertion (b) can be broken down into several assumptions:

> (v)   there is a divine creator and a set of creatures, qualitatively different from each other, and

> (vi)   it is morally wrong to cross the boundary between them.

In (b'), assumption (v) has been discarded – there is no reference to a qualitatively different creator and creatures. New assumptions have been introduced, namely that

> (vii)   the reciprocal relationship among human beings must be respected, and

> (viii)   the principle of equal freedom, as specified in (a'), must be respected.

Assumption (vi), that there is a boundary which one wrongly crosses by mixing the roles of creature and creator, remains. Again, the transition from (b) to (b') follows the pattern of a type two ethical appropriation. Assumption (v) is the original substance of piety surrounding the ethical insight of (vi), which after appropriation is clad in a new substance of piety made up of (vii) and (viii). The reason for the wrongness of mixing creature and creator has changed – once it was because of the morally charged qualitative difference between God and his creation, now it is because doing so fails to respect the principle of equal freedom or the reciprocal relationship among humans – but the existence of the boundary, and the wrongness of crossing it, remain. Habermas uses this ethical insight to argue against the genetic modification of human beings, by human beings. (FHN 114-5)

Habermas' own example of sacred-to-secular translation consists of a pair of type two ethical appropriations. When he discusses the translations made by other theorists, he typically cites what I call conceptual appropriations. In the same passage in which he first introduces translation, he refers to Kant's use of Christian motifs. Kant, we are told, was the first postmetaphysical thinker to carry out "a critical *assimilation* of religious contents," transforming the notion of divine command into moral duty. Habermas adds that his "further attempt to

translate the notion of 'radical evil' from biblical language into the language of rational religion may seem less convincing." (FHN 110) Habermas seems to be equating Kant's use of these motifs with religious citizens' use of "God created man in his own image."[71] In "Prepolitical Foundations of the Constitutional State?," Habermas explains that postmetaphysical philosophy can borrow and transform religious concepts without deflating them: "The translation of the theological doctrine of creation in God's image into the idea of the equal and unconditional dignity of all human beings constitutes one such conserving translation. It makes the content of biblical concepts available to the general public of unbelievers and members of other faiths beyond the boundaries of a particular religious community." (BNR 110) Here Habermas uses Walter Benjamin as an example; elsewhere, he describes Benjamin's idea of anamnestic solidarity as a translation/conceptual appropriation of the Last Judgement.[72] He juxtaposes this passage with a description of the postsecular model, including the requirement that non-religious citizens "take part in the efforts to translate relevant contributions from religious language into a publicly intelligible language." (BNR 113) There is no reason to think that he gives a different meaning to the word "translate" in these neighbouring passages.

In "Religion in the Public Sphere," the text which deals most directly with the postsecular model, Habermas seems to regard the translation of religious statements carried out by citizens as an extension of the appropriation of metaphysical motifs carried out by philosophers. He says that he is discussing the "ambivalent attitude of postmetaphysical thinking to religion because it also expresses the cognitive presupposition for the willingness to cooperate of secular citizens." Both processes, then, draw on the same attitudes or cognitive adjustments. "This ambivalent attitude to religion corresponds exactly to the epistemic attitude that secular citizens

---

[71] Habermas returns to Kant at length in "The Boundary Between Faith and Knowledge: On the Reception and Contemporary Importance of Kant's Philosophy of Religion" (BNR 209-247), where he discusses Kant's conceptual and ethical appropriations. He says that Kant had the "goal of redeeming religious contents through translation." (BNR 228)

[72] See BNR 241. Habermas assumes, by implication, that Benjamin's philosophy is more accessible to the general public than the religious concepts of the Last Judgement or the Messiah.

must adopt if they are to be prepared to learn something from the contributions of their religious counterparts to public debates which are potentially translatable into a generally accessible language." (BNR 143)

The most reasonable interpretation of Habermas' comments is that he regards sacred-to-secular translation and postmetaphysical appropriation as identical. His one concrete example of a translation appears to be, by my definitions, an ethical appropriation; the examples he mentions in passing are typically conceptual appropriations. Aside from the familiar cases of Kant, Hegel, Marx, Benjamin and Adorno, Habermas also mentions Kierkegaard, Bloch, Levinas and Derrida, without specifying which of their concepts are translations or appropriations. Interestingly, he says that Heidegger's "remembrance of being" is "a veiled borrowing, not a translation,"[73] again suggesting that translated or appropriated concepts undergo considerable change – they are not simply transplanted in one piece into a new register. (T 309, FHN 110-1) It seems clear that to Habermas, what reflexive non-believing citizens do with religious statements in the informal public sphere is the same as what postmetaphysical thinkers do with metaphysical motifs, or with ethical values.

But Habermas' equation of translation with appropriation is questionable. I will argue, firstly, that the two procedures should be distinguished from each other, and secondly, that translation cannot work in the same way as appropriation. I am not alone in making this first assertion. Simone Chambers points out that the "use of translation in this context [of appropriation] is somewhat different from the 'institutional translation proviso'."[74] She notes that appropriation is a process carried out by philosophers, rather than ordinary citizens.[75] Max

---

[73] "Reply to my Critics," in *Habermas and Religion*, p 387.

[74] Simone Chambers, "How Religion Speaks to the Agnostic: Habermas on the Persistent Value of Religion" *Constellations* 14:2 (2007), p 216.

[75] In context, Chambers is talking about how the burden of translation falls on religious citizens in unmodified, secular deliberative democracy. She goes along with Habermas' argument that it is a burden which falls exclusively on believers.

Pensky agrees with Chambers that there are two procedures at work in Habermas' writings, one carried out by philosophers, the other by citizens in the public sphere, although he refers to them both as "translation."[76] Darren Walhof points out that "Habermas tends to equate the mining of religion's resources with the 'translation' of religious reasons into secular reasons," and adds that it does not seem quite right to call such a process translation.[77]

Nicholas Wolterstorff, who has proved to be quite critical of the use Habermas makes of his arguments, states the distinction most clearly. He argues that what postmetaphysical philosophers do with religious concepts is better described as appropriation than translation, and goes on to differentiate between the two processes:

> We have all had the experience, upon listening to someone of a quite different persuasion from our own, of seeing the reality that he was trying to get at even though we ourselves would never put it that way. Though we dissent from the position he affirms, we see what he was trying to get at. We then put this in our own words; we *appropriate* it. We don't translate *what he said* into a different language; we appropriate *what he was trying to get at*. Habermas's thought is that appropriation, so understood, is what the postmetaphysical philosopher mainly aims at in his dialogue with religion.[78]

Like Chambers and Pensky, Wolterstorff makes the point that conceptual appropriation is a process carried out by philosophers, for philosophers. Its goal is to furnish postmetaphysical philosophy with a fund of conceptual motifs, not to allow religious citizens to participate in the public sphere. All these theorists suggest that whatever reflexive citizens are doing with religious inputs in Habermas' model, it does not seem to be the same as what postmetaphysical philosophers do with metaphysical concepts, or postmetaphysical thinkers in general do with ethical values.

---

[76] Max Pensky, "Solidarity with the Past and the Work of Translation: Reflections on Memory Politics and the Postsecular," in *Habermas and Religion,* pp 303-4.

[77] Walhof op. cit., pp 233-4.

[78] Nicholas Wolterstorff, "An Engagement with Jürgen Habermas on Postmetaphysical Philosophy, Religion, and Political Dialogue," in *Habermas and Religion*, pp 99-100.

So my first assertion, that translation and appropriation are two different processes, is not new. In the remainder of this section, I will argue for my second assertion: that sacred-to-secular translation in the postsecular model *cannot* function in the same way as postmetaphysical appropriation. I have identified three varieties of appropriation. The acknowledgement requirement of the postsecular model rules out two varieties of the procedure, and the nature of the formal public sphere rules out the third.

Let us first consider conceptual appropriation. On a mundane level, the two processes display many differences. As Chambers, Pensky and Wolterstorff observe, conceptual appropriation is a process carried out in the realm of theory, by philosophers. Its purpose, at least in Habermas' writings from the 1980s and 1990s, is to make concepts from earlier philosophical paradigms available to later thinkers. As such the integrity or recognisability of the concepts are not at issue – the bare motif of a concept is extracted, and its "substance of piety" is discarded. Sacred-to-secular translation is carried out in the informal public sphere, by reflexive citizens. It has more than one purpose – it balances the cognitive burdens of believers and non-believers, and allows "monoglot" religious citizens to see themselves as the authors and addressees of legitimate laws. This last goal, in the context of the discourse theory of law and democracy, means that religious citizens must be able to acknowledge translations of their contributions as they pass into the formal public sphere. It follows that secular translations must be similar enough to their religious source statements that monoglot believers can acknowledge them. Prima facie, then, the two processes appear to be carried out by different people, in different contexts, for different goals.

I do not suggest that these differences are decisive: a single procedure may be used for different purposes, just as a single medicine may treat several diseases. Neither would I want to suggest that philosophers and ordinary citizens inhabit different universes. Philosophers are citizens, and anyone may become a philosopher. But there is a question of the plausibility of

equating conceptual appropriation with sacred-to-secular translation. Does Habermas expect ordinary citizens – however reflexive – to perform the same philosophical feats as Benjamin or Adorno, as part of their daily business? One of his arguments against strict secularism is that it is too cognitively demanding of religious citizens, and that the liberal state should not ask them to perform a task which they find "impossible." Is it reasonable to ask citizens to do with religious contributions what Hegel did with Böhme's theodicy? Most importantly, a conceptual appropriation would never meet the acknowledgement requirement. What is produced at the end of the procedure is the bare motif of the appropriated idea, rebuilt out of new conceptual materials. To expect religious citizens to acknowledge such radically transformed versions of their inputs as their own is unreasonable. As Austin Harrington points out, Habermas himself has expressed doubt on this point.[79] In the "Transcendence" essay, he talks about the "translations" of religious concepts carried out by critical theologians (T 310), noting that they diverge to a great extent from the beliefs of ordinary adherents. Habermas discusses the methodically atheist theology of Jens Glebe-Möller, and questions whether it leaves the religious language-game intact. "But I ask myself *who* recognizes himself in [Glebe-Möller's] interpretation," (T 311) he adds. What goes for these conceptual appropriations also goes for sacred-to-secular translation, on this interpretation.

The case of ethical appropriation is more complex. Any postmetaphysical thinker may carry out an ethical appropriation – the procedure is not limited to philosophers. There is no reason that the procedure should not be carried out in the public sphere. The goals are different – ethical appropriation aims at supplying postmetaphysical thinkers with values and insights for use in their ethical discourses – but the same method could be used for both. The difficulty with equating sacred-to-secular translation with ethical appropriation arises from the nature of the procedure.

---

[79] Harrington op. cit., p 552.

Consider type one ethical appropriation. If sacred-to-secular translation worked in this way, then the substance of piety surrounding a religious input would have to be taken over, whole cloth, into the language of the formal public sphere. This would not be a "secular translation," in the sense of a statement being transferred from a religious source language to a secular target language; if the target language contains all the terms needed to state the original substance of piety of the religious statement, it would not be *secular* language.[80] To put it in more plainly, a sacred-to-secular translation which results in a substantially religious statement is a failure, not only as a translation, but according to the requirements of Habermas' model. Habermas wants the formal public sphere to remain strictly secular. This is the point of the ITP and the prohibition on religious arguments in courts and parliaments. If a translation in the form of a type one ethical appropriation allows not only (formerly) religious insights, but *chunks of the religious worldviews in which they originate* to pass through into the formal public sphere, then it has failed.

Consider type two ethical appropriation. This seems like the better option. As I have argued, Habermas' sole example of a sacred-to-secular translation appears to follow the pattern of this kind of appropriation. The resulting statement would consist of an ethical insight plucked out of its religious context and enclosed in a new, secular substance of piety. Thus it would preserve the secular nature of the formal public sphere. Habermas assumes that there is a gulf of understanding between religious and non-religious citizens, so it seems likely that a type two appropriation/translation would be necessary. For just this reason, however, this interpretation of translation falls victim to what I earlier called the acknowledgement problem. It is an implication of Habermas' conception of deliberative democracy, and its postsecular modification, that secular translations of religious inputs must meet the acknowledgement

---

[80] I will return to this point in the next section.

requirement. The religious citizens who first made them must be able to acknowledge them, whether as inputs to the informal public sphere or as laws produced by the formal public sphere.

Now, would religious citizens be able to acknowledge translations of their inputs as their own, if "translation" means type two ethical appropriation? This depends on how religious citizens are characterised. And Habermas is inconsistent on this point. As I have noted, he sometimes takes Weithman's position and imagines that they are unable to feel any "pull" at all from secular reasons and arguments. At other times, he goes along with Wolterstorff and assumes they while they can appreciate the force of secular reasons, these reasons do not feel authentic to them, and they should not be obliged to make use of them in public. It is unlikely that Weithman's believers could acknowledge type two ethical appropriations of their religious inputs. If secular arguments leave them cold, then so would secular re-presentations of the ethical values contained in their religious statements. Acknowledgement is not simply a matter of understanding. I am assuming that a religious believer as Weithman (and, on occasion, Habermas) characterises them would look at a secular version of one of their inputs and say, "You have removed the most important and meaningful part of what I was saying. This is no translation at all – it is a misrepresentation."[81] Wolterstorff's believers would be more attuned to these kinds of translations of their inputs, in the sense that they could appreciate the force of the secular arguments which their contributions have become. But even this might not be enough for acknowledgement, and thus legitimation, to take place. Feeling the force of the reasons contained in a target text is not the same as acknowledging them to be *versions of the same reasons* found in the source text. Wolterstorff's believers, reflexive as they are, might still

---

[81] Andrea Baumeister makes exactly this point: "Even if a Habermasian translation programme succeeded in transposing the fundamental moral intuitions… into a generally accessible secular language, it could not capture the special significance or weight that attaches to these reasons from a religious perspective. While such a secular translation may well be accessible to religious citizens, it will from a religious perspective miss the most pertinent reasons." See Andrea Baumeister, "The Use of Public Reason by Religious and Secular Citizens: Limitations of Habermas' Conception of Religion in the Public Realm," *Constellations* 18:2 (June 2011), p 229.

feel that something irreplaceable is lost when the religious substance of piety enclosing the ethical insight is discarded. Cooke too worries that Habermas "passes too easily over the difficulties involved in transposing images, figures and narratives from a metaphysical framework into a postmetaphysical one, simply taking for granted that the transposed contents will retain a semantic power."[82] She is far from alone in making this criticism. Type two ethical appropriation may be the best candidate for sacred-to-secular translation, but it is hard to believe that the kind of religious believers for whom Habermas builds his model could acknowledge its products.

Defenders of Habermas might argue that I am overstating the difficulties which believers would experience in acknowledging appropriated statements. Habermas does say that a religious citizen must have "the epistemic ability to consider one's own religious convictions reflexively from the outside and to connect them with secular views." (BNR 130) What he means by this ability to "connect" is not clear, but it could be interpreted in terms of acknowledgement. It is possible, of course, to imagine religious citizens who are reflexive enough to acknowledge ethical appropriations as translations of their religious inputs.[83] It could be argued that any believers who had undergone the three cognitive adjustments described in section 3.7 would be capable of this. Consider, however, what this would entail, if translation is understood as type two ethical appropriation. In order to acknowledge translations of their inputs, religious citizens would have to accept that it is the ethical insights contained in religious statements which matter, not the religious nature of the statements, and they would have to be capable of mentally separating these insights from their substance of piety. This would require an extremely high level of reflexivity about one's own beliefs. Now, why would a religious citizen this reflexive have any objections to strict secularism? If the ethical insights are what matter,

---

[82] Cooke (2006), p 201.

[83] It is also possible to imagine religious citizens who are philosophically literate enough to recognise conceptual appropriations, and non-religious citizens skilful enough to carry them out, but I do not want to stretch credibility.

not the vocabulary in which they are (contingently) embedded, why object to using a particular vocabulary? Such citizens would be unlikely to suffer from identity-splitting, and the need to "translate" their own contributions would not trouble them – they would be polyglots, not monoglots. My point is that the kind of religious citizen capable of acknowledging type two ethical appropriations would be very remote from the kind Habermas builds his model around. They would not find secularism onerous in the first place, and would feel no need for postsecularism.

The problem of transformation and acknowledgement afflicts any attempt to equate sacred-to-secular translation with appropriation. Translation must result in secular statements which can be acknowledged by those who made the original religious statements – Habermas' postsecular model requires it, for legitimation to take place. But appropriation, as Habermas himself states, drastically transforms the nature of the appropriated content. A transformed statement may or may not be capable of being acknowledged by the person who made the original – this depends on how radical the transformation is, and the person's powers of acknowledgement. Habermas' characterisation of religious believers, however, tips the scale towards a failure of acknowledgement. In order to justify his postsecular model, he has to emphasise the distance in understanding between believers and non-believers; but if this model is to produce legitimate laws and policies for believers, they must be able to acknowledge secular translations of their contributions as their own. No variety of appropriation will perform this task *for religious believers as Habermas imagines them*. Contrary to Cooke's interpretation, and to many of Habermas' own comments, sacred-to-secular translation cannot be a species of appropriation. In the following section I will consider what else it might be.

3.10 Pyrrhic Translations

If translation cannot be appropriation, what else might it be? In this final section I will assess the procedure as if it was a version of interlinguistic translation, with the religious register as the source language and the secular register as the target language. I will conclude that this approach to sacred-to-secular translation is just as problematic as Habermas' equation of translation with appropriation.

It is not my wish to argue that religious statements are *untranslatable*. Incommunicability, in Primo Levi's words, is a linguistic horror.[84] Paul Ricoeur argued that rather than thinking in terms of translatability and untranslatability, we should assess translations in terms of their faithfulness to, or betrayal of, the source text.[85] My thesis will be that secular translations of religious texts cannot help but betray their source texts, to such an extent that they cannot perform the function Habermas ascribes for them in his postsecular model. I call these *pyrrhic translations*. A translation of a statement is pyrrhic when it loses or destroys most of the statement's essential content – the humour of a joke, the truth-value of a factual assertion, the beauty of a line of poetry. I will outline two ways in which sacred-to-secular translations cannot help but be pyrrhic. Habermas himself acknowledges the extreme difficulty of rendering religious statements into secular language. Religion, he says, has an "opaque core" which "remains as profoundly alien to discursive thought as the hermetic core of aesthetic experience." (BNR 143)[86] Commentators on his work have made similar points.[87] I simply want

---

[84] Primo Levi, *The Drowned and the Saved* (London: Abacus, 2003), p 68.

[85] Paul Ricoeur, "On the Paradigm of Translation," in *On Translation* (London and New York: Routledge, 2006), p 14.

[86] See also BNR 129-30, 132 n 37, 242-3.

[87] See Cooke (2006) pp 201-2 Cooke (2011)p 485, Cooke in HR p 260, Ennis (2007) p 882, p 890, Ennis (2010) p 585, Lafont (2007) p 245, Lafont (2009) pp 134-5, Harrington (2007) p 552, Ungureanu (2008) p 408 n 12, p 417 n 46, Boettcher (2009), Ricken in AWM pp 57-8, Richard J. Bernstein (2010) p 162, Carlo Invernizzi Accetti, "Can Democracy Emancipate Itself from Political Theology? Habermas and Lefort on the Permanence of the Theologico-Political" *Constellations* 17:2 (2010) p 259, Baumeister (2011) p 229,

to give a rigorous account of this difficulty. If my assessment is correct, then Habermas' model of postsecular deliberative democracy, which relies on sacred-to-secular translation, must be rejected.

Let us consider Habermas' procedure from the perspective of translation theory. The *skopos theory* of translation posits that the function or purpose (*skopos*) of the target text must be the governing factor in the translation process. Translators must be orientated by the role which their text, in its target language version, is meant to play. The quality of a translation can only be assessed with its skopos in mind.[88] Of course, a translation may have multiple skopoi. Fidelity to the source language original is a frequent, but not inevitable skopos. What are the skopoi of sacred-to-secular translations in Habermas' model? Balancing the cognitive burdens of religious and non-religious citizens, and allowing the former to participate in the legitimating process of contributing to the public sphere, are the obvious ones. The acknowledgement requirement also gives rise to a skopos: religious citizens must be capable of acknowledging the target language versions of their contributions as their own. Finally, Habermas' insistence on the strictly secular nature of the formal public sphere gives us a fourth skopos: the integrity of the target language must not be violated.

It is easy to see that these last two skopoi pull in different directions, given Habermas' characterisation of religious believers. It could be said that sacred-to-secular translations have contradictory skopoi – they must be capable of being acknowledged by citizens who either cannot understand or feel no "pull" from secular reasons and arguments, and at the same time they must remain within the boundaries of secular language. This is another aspect of the problem which prevents type one ethical appropriations from working as translations. As I noted

---

pp 231-2, Bernd Irlenborn, "Religion in the Public Sphere: Habermas on the Role of Christian Faith" *The Heythrop Journal* 53:3 (2012) p 437, Aguirre (2013) pp 643-4, Pensky op. cit., pp 319-20.

[88] See Christina Schäffner, "*Skopos* Theory," in Mona Baker ed., *The Routledge Encyclopaedia of Translation Studies* (London and New York: Routledge, 2001), pp 235-8. The theory was first formulated by Hans J. Vermeer and Katarina Reiss.

in the previous section, a type one ethical appropriation would easily meet the acknowledgement requirement, since in this type of appropriation an ethical insight is transposed into postmetaphysical language along with its surrounding substance of piety. However, it would completely violate the skopos of maintaining the integrity of secular language in the formal public sphere. It seems doubtful that any translator, however skilful, could juggle these contradictory skopoi. A fortiori, asking ordinary non-religious citizens to do so seems unreasonable.

As Ricoeur observes, it is easy to argue that translation is impossible, but the fact remains that it does take place.[89] The problem of contradictory skopoi need not be insurmountable. Theorists generally agree that translation is most difficult when the source and target languages are very different from each other, but even in these cases some form of translation is usually possible. Religious and secular language are, on Habermas' account, profoundly different. Recall Habermas' account of religious language. As I noted in Part 1, in his work on social theory from the 1970s and 1980s Habermas sees religious language as undifferentiated – it is a "syndrome" of entangled validity claims, attitudes and world-concepts. In contrast to modern discourse, dimensions of validity are not separable and individual validity claims cannot be thematised. One might expect Habermas to abandon this theory after moving in a postsecular direction, but in fact he shows every sign of retaining it. As recently as his "Reply" in *Habermas and Religion*, he asserts that, from an agnostic perspective[90], "religious 'truths' are formulated in concepts that are prior to the usual differentiation into descriptive, evaluative, and normative statements."[91] This would explain his remarks about the "discursive

---

[89] Paul Ricoeur, "A 'Passage': Translating the Untranslatable" in Ricoeur op. cit., p 32.

[90] Which presumably includes his own, postmetaphysical perspective.

[91] Habermas, "Reply to My Critics," *Habermas and Religion* p 367.

extraterritoriality" of the core of religious experience. (BNR 130) Thomas McCarthy certainly believes that Habermas is still working with this account of religious language.[92]

Let us assume that this is the case. The difficulty of sacred-to-secular translation, then, can be ascribed to the divergence between the sacred source language and the secular target language: the former is an undifferentiated "syndrome of validity," while the latter has undergone a tripartite division of worlds, attitudes and validity claims. The linguist Roman Jakobson famously identified some of the problems which translators encounter when they try to bridge the gap between structurally different languages. Firstly, the source language may contain terms which do not exist in the target language. This is certainly the case with sacred-to-secular translation – religious language contains religious terms, and secular language by definition does not. Jakobson uses the example of translating a text containing the word "cheese" into the language of a culture which does not consume cheese. He does not think that this is a serious obstacle to translation. "Whenever there is deficiency," says Jakobson, "terminology may be qualified and amplified by loanwords or loan-translations, neologisms or semantic shifts, and finally, by circumlocutions… No lack of grammatical device in the language translated into makes impossible a literal translation of the entire conceptual information contained in the original."[93] The problem of translating "cheese" vanishes if the translator adds a term or phrase meaning "food made of pressed curds" to their target text.[94] The target

---

[92] McCarthy, "The Burdens of Modernized Faith and Postmetaphysical Reason in Habermas's 'Unfinished Project of Modernity," Habermas and Religion p 119.

[93] Roman Jakobson, "On Linguistic Aspects of Translation," in Rainer Schulte and John Biguenet eds., *Theories of Translation: An Anthology of Essays from Dryden to Derrida* (Chicago and London: University of Chicago Press, 1992), p 147.

[94] Ibid, p 144.

language can be expanded by adding new terms to it, and in this way translation can enrich a language.[95]

Of course, this elegant solution is not available in the case of sacred-to-secular translation. It would violate the skopos of preserving the integrity of secular language. Adding religious neologisms to secular language would mean that it is no longer secular; this is why type one ethical appropriation is not acceptable as translation. If sacred-to-secular translation allowed explicitly religious reasons and arguments to enter the formal public sphere, then one of the fundamental constraints of postsecular deliberative democracy would have been violated. Habermas' model would become identical to Weithman's or Wolterstorff's.

In these examples, the source language contains something which is missing from the target language. Jakobson suggests that translation encounters more serious problems when the reverse is true – when the target language draws a distinction which the source language does not, or has greater informational requirements. For example, English does not distinguish between dual and plural terms: the same word is used whether there are two or a thousand of something. Other languages do draw this distinction. If the English statement "She has brothers" is translated into such a language, says Jakobson, the translator is faced with a choice between

---

[95] This is a point Schleiermacher makes in his influential essay on translation. He divides translations into two types: "Either the translator leaves the writer alone as much as possible and moves the reader towards the writer, or he leaves the reader alone as much as possible and moves the writer towards the reader." This is usually referred to as the distinction between *domesticating* translations – which make the source language sound like the target language – and *foreignising* translations – which make the target language sound like the source language. See Friedrich Schleiermacher, "On the Differing Methods of Translation," in Schulte and Biguenet op. cit., p 42. Schleiermacher, like many of his contemporaries, strongly favoured foreignising translations, partly because they can enrich the target language with new words and concepts. The quandary of sacred-to-secular translation in Habermas' model can be expressed in Schleiermacher's terms. Domesticating translations fulfil the skopos of preserving the integrity of secular language, but do not meet the acknowledgement requirement (given Habermas' account of religious believers); foreignising translations meet the skopos of acknowledgement, but violate the integrity of secular language. It might make sense to describe conceptual and ethical appropriation as types of foreignising translation, since the purpose of such appropriation is to expand and enrich postmetaphysical discourse. Habermas talks about Kant's appropriations of Christian concepts in these terms – see BNR 224-6. But sacred-to-secular translation is another matter.

saying "She has (*two*) brothers" and "She has (*more than two*) brothers." The original statement

gives them no guidance – the target language is asking for something which the source language

does not supply. [96] This is also the case for sacred-to-secular translations. Secular language is a

subset of differentiated language – it distinguishes between claims of expressive, normative and

factual validity, and the accompanying attitudes and world-relations. Religious language does

not. Jakobson believes that this problem can also be solved, albeit with more difficulty than the

problem of missing terms in the target language. Different languages, he argues, make different

informational requirements at the grammatical level, but at the cognitive level any language can

convey any content, if need be by resorting to neologisms and paraphrasing. "Languages differ

in what they *must* convey and not in what they *may* convey."[97] In his example, the translator

could find out from other sources how many brothers the person concerned has, and choose

the appropriate term, or they could use a paraphrase such as "She has either two or more than

two brothers." Although he is adamant that any "assumption of ineffable or untranslatable

cognitive data would be a contradiction in terms," Jakobson adds that "in jest, in dreams, in

magic, briefly, in what one would call everyday verbal mythology and in poetry above all, the

grammatical categories carry a high semantic import. In these conditions, the question of

translation becomes much more entangled and controversial."

This is why Jakobson's solution to this second problem is not applicable to sacred-to-

secular translation. The "extra information" required to meet the needs of the secular target

language is the kind of information which by definition cannot be found in the sacred source

language. The grammatical structure and semantic content of religious language are, according

to Habermas, mutually dependent (a point I will return to below). The target language requires

differentiated validity claims, attitudes and world-relations, but the source language only *exists*

---

[96] Jakobson op. cit., p 148.

[97] Ibid, p 149.

by virtue of these elements *not* being differentiated. To disentangle them would be to destroy religious language and the things it can say, not to translate it. (T 310) Needless to say, a translation which attempted to do this would violate the skopos of acknowledgement.

Jakobson identifies the obstacles to translation created by the mismatch between structurally different languages, and shows that they are not insurmountable. His solutions rely on the plasticity of languages. Living, natural languages are infinitely adaptable; there is no limit to what they can say.[98] But religious and secular "languages" are not living, natural languages. Secular language is an artificially limited subset of any given language, stripped of religious terms. It is defined by what it does not, cannot, say. Religious language, on Habermas' account, is an epiphenomenon of undifferentiated attitudes, world-relations and validity claims. It is adaptable up to a point, but it cannot differentiate its entangled elements: to do so would be to destroy itself through linguistification. Sacred-to-secular translation is in the unfortunate position of suffering from the problems Jakobson identifies, due to the structural difference between its source and target languages, but not being able to make use of his solutions, due to the nature of the languages involved. The situation is made worse by the contradictory skopoi of Habermas' translation procedure: sacred-to-secular translations must be recognisable to religious citizens as translations of their inputs, and at the same time must not violate the integrity of secular language.

These considerations might be enough, on their own, to demonstrate that sacred-to-secular translations must be pyrrhic. However, I would like to make a final argument to the effect that translations of religious language, as it is understood by Habermas, are always pyrrhic. I will begin with a broad definition of translation: statements have a *form* and a *content*, and translation is a process of *transposing this content into a new form*. I am using the term "content" loosely. It might be the meaning of the statement, the information it conveys, the

---

[98] Jakobson op. cit., p 147.

impression it creates on recipients, and so on. The form in question is the language in which this content is encoded. It follows that when the form and content of a statement are easily separable, translation will go smoothly. The content of the statement "it is raining today" is easily separable from its form, and as a result a translator will experience no difficulty in putting that content into a new form: "mae hi'n bwrw glaw heddiw," "es regnet heute," and so on. Translation runs into difficulty, in contrast, when content "sticks" to form. If the content of a statement somehow relies on the form of the statement for its meaning, truthfulness, or some other quality, if the two are entangled, then attempts to transpose that content into another form are likely to produce pyrrhic translations.

Some examples will clarify my argument. The most striking case of form/content dependence is found in self-referential statements. Tyler Burge points out that the English sentence

(A) This very sentence begins with a four-letter demonstrative

if translated literally into German, becomes

(B) Dieser Satz fängt mit einem hinweisenden Artikel mit vier Buchstaben an

which loses the truth-value of (A). It would be better, for some purposes, to translate (A) as

(C) Dieser Satz fängt mit einem hinweisenden Artikel mit sechs Buchstaben an

which has the same truth-value as the original, but at the same time has a different cognitive content.[99] (A) and (C) are not the same – "sechs" is not a translation of "four." In this case a literal translation of the source language statement produces a target language statement which

---

[99] Tyler Burge, "Self-Reference and Translation," in F. Guenther and M. Guenther-Reutter eds., *Meaning and Translation – Philosophical and Linguistic Approaches* (London: Gerald Duckworth, 1978), pp 138-9. Burge is not making the same argument I am, about translations of statements where content is dependent on form being pyrrhic, but his examples illustrate my point.

lacks one of the most important features of the original. The translator can instead produce a

statement in the target language which replicates this feature of the source language original,

but this is a *different statement*, saying a different thing. I would describe both these cases as

pyrrhic translations. Burge gives another example, where the sentence

(D) The seventh word of this sentence has three letters

can be translated so as to maintain its truth-value either as

(E) Das siebte Wort dieses Satzes hat vier Buchstaben

or as

(F) Das sechste Wort dieses Satzes hat drei Buchstaben.[100]

Something fundamental goes missing in the process of translating (D), namely the statement's

truth-value. Both (E) and (F) have truth-value, but they state different facts to (D). In both cases,

the target language versions of the statement say something different to the original. If these

are translations, they are pyrrhic. What Burge's examples bring out is the issue of form/content

dependence. In a self-referential statement, the form of the statement is part of its content.

When the content of a statement adheres its form in this way, the process of putting that

content into a new form – of translation – becomes pyrrhic.

Poetry and wordplay, which are notoriously difficult to translate, are also examples of

form/content dependence. Puns, for example, rely on homonymy and homophony among the

words of a particular language, as well as the associations and connotations of those words. The

content of a pun, its humorous effect, relies heavily on the form of the language in which it is

made. An English pun on the homophones "lead" and "led" would not work in Welsh, where the

words for the metal and the past participle are "plwm" and "arwain," respectively. A literal

---

[100] Ibid, p 142.

translation of the pun would lose the essence of its content, its humorous effect. Such a translation would be pyrrhic. Translators of texts which contain puns and other forms of wordplay have long acknowledged this difficulty. They usually respond to source language puns by abandoning them and producing entirely new puns in the target language.[101] It is not clear whether this procedure should be called "translation" at all.

James Joyce's novel *Finnegans Wake*, one of the most dramatic examples of extended wordplay ever produced, is a case in point. It has of course been translated, but the consensus among Joyce scholars is that these translations are usually pyrrhic. Leo Knuth, summing up the view of a panel of translators, says that in the language of *Finnegans Wake* "the *inextricable relationship* between sound and sense (or nonsense), between music and meaning (a particularly slippery concept), makes the task [of translation] well-nigh impossible (the majority opinion) or at least much harder than the translation of other books (the opinion of the most sanguine minority)."[102] In other words, a close connection between form and content makes translation pyrrhic. Some commentators take a different view. W.V. Costanzo argues that Joyce's instances of wordplay are often translatable, in fairly literal terms. In André du Bouchet's French translation, for example, "for to watsch," becomes "pour se lavoir." The English pun on watch/wash is exactly replicated by the French la voir/laver.[103] Costanzo admits, however, that only a minority of Joyce's puns can be translated so directly. Even in the translation which Joyce himself oversaw, the French puns differ significantly from the English ones. A reference to "this

---

[101] See Keith Harvey, "Compensation" in *Routledge Encyclopaedia*, pp 37-40. Peter Alan Low insists that puns are translatable – see his "Translating Jokes and Puns," *Perspectives* 19:1 (2011). Low uses a loose definition of "translation," however, which resembles connotative rather than denotative translation (see n 105 below). In practice, most of his examples are new coinages in the target language, and he admits (p 63) that less than 1% of puns are fully translatable.

[102] Leo Knuth, "The *Finnegans Wake* Translation Panel at Trieste" *James Joyce Quarterly* 9:2 (1972), p 267. Emphasis added.

[103] W.V. Costanzo, "The French Version of 'Finnegans Wake': Translation, Adaptation, Recreation" *James Joyce Quarterly* 9:2 (1972), p 228.

man of hod, cement and edifices" becomes "cet home de sac, de bêton et de bâtises." Most of

the double meanings in French, says Costanzo, "have little to do with the original. A man of odd

jobs, for example ('home de sac et de baton,' 'bêton' = concrete) is hardly the same as a man of

God, and a worker who makes jokes ('bêtises') is quite different from one who simply raises

edifices."[104] He observes that at times, "the resistance to translation is so great that du Bouchet

omits entire sentences. This is especially true of the passages with special rhythmic and

alliterative effects." Costanzo attributes this divergence in the content of wordplay to a

structural difference between the source and target language – compared to French, English has

a greater number of onomatopoeic words and variations in its rhythm.[105] He concludes that "For

Joyce, translating *Finnegans Wake* is a matter of rewriting the book."[106] Once again, if this is

translation, it is pyrrhic. As Knuth observes, "When a pun in the source language is replaced by

a different pun in the target language, can the translation still be said to reflect the original?"[107]

---

[104] Ibid, p 230.

[105] Ibid, p 231.

[106] Ibid, p 229

[107] Knuth op. cit., pp 268-9. Such a procedure can be called translation if one adopts a connotative, rather than denotative definition. Denotative translation transfers the *cognitive content* of a statement from source to target language, connotative translation replicates the *effect* produced on source language receivers in target language receivers. Its aim is to make speakers of the target language have the same reaction to the translated text as speakers of the source language did to the original. See Werner Koller, "Equivalence in Translation Theory," in Andrew Chesterman ed., *Readings in Translation Theory* (Helsinki: Oy Finn Lectura Ab, 1989). This is Low's approach to the translation of jokes and puns – all that matters is that target language speakers experience humour; the contents of translated statements are irrelevant, so creating a new pun which causes them to experience humour is acceptable. Habermas' example of sacred-to-secular translation could, at a stretch, be interpreted in this way. Turning "God created man in his own image" into "a gamete fertilized ex utreo is a subject of human rights" is a successful connotative translation, *if* the skopos of sacred-to-secular translation is to ensure that non-religious citizens have the same reaction to the target language statement as religious citizens have to the original. *Human rights* "translates" *image of God* in the sense that both phrases provoke a reaction of respect, normative weight, and inviolability. But there is no sign, anywhere in Habermas' writings, that producing similar emotional reactions in different groups of citizens is part of the skopos of sacred-to-secular translation. There is every sign that sacred-to-secular translations are meant to result in secular statements with the same cognitive content as religious statements. A connotative translation, in addition, does not attempt to meet the acknowledgement requirement.

My argument, in summary, is that any translation of a statement which displays form/content dependence will be pyrrhic, when the source and target languages are structurally dissimilar. Translation is a process of putting the content of a statement into a new form; when content is dependent on form, this procedure becomes problematic. Now, do religious statements display form/content dependence? According to Habermas, they do. They are deeply dependent on the structure of the language in which they are made. He thinks that the content of statements about "the creator and redeemer God… theodicy, [and] the event of salvation" rely on the form of religious language, namely that "the ontic, normative, and expressive aspects of validity… remain fused together." (T 310) In an analogous way, the truth-value of the statement "Dieser Satz fängt mit einem hinweisenden Artikel mit sechs Buchstaben an" depends on the form of the German language, namely that the word "dieser" has six letters. If these religious concepts can *only* be expressed in undifferentiated language, then the attempt to translate them into differentiated language will be problematic in the extreme. Secular translations of religious statements will always be pyrrhic.

Would a pyrrhic translation meet the acknowledgement requirement? It seems clear that it would not. A pyrrhic translation of a factual assertion destroys its truth value; a pyrrhic translation of a line of poetry destroys its beauty; a pyrrhic translation of a joke or piece of wordplay destroys its humour; and a pyrrhic translation of a religious statement destroys the very element which matters most to believers – its sacred meaning. They would not acknowledge such bowdlerisations of their reasons and arguments as their own. Sacred-to-secular translation cannot perform the function which Habermas ascribes to it. The procedure fails, and with it Habermas' model of postsecular deliberative democracy.

Conclusion

We do not change secular questions into theological ones. We change theological questions into secular ones.

Karl Marx, *On the Jewish Question*

Deus Ex Machina

In Ancient Greek theatre, the climax of a play would sometimes be interrupted by the appearance of a god on stage. Descending from a crane or rising from a trapdoor, the actor representing the god could rescue the human characters from their predicaments or resolve an insoluble situation. The plot of Euripides' *Alcestis* concerns the eponymous wife of King Admetus, who must die in place of her husband. At the end of the play, Euripides abruptly introduces Heracles, who retrieves Alcestis from death after she sacrifices herself.

[1] The *deus ex machina*, the god from the machine, allowed playwrights to break their own rules, and thus to extract themselves from impossible situations.

Philosophers sometimes do the same. By the end of the 1990s, Habermas was faced with an insoluble problem, largely of his own making. The spread of hard naturalism and the development of biotechnology threatened to disrupt the anthropic basis of his philosophical project, and much else besides. But the postmetaphysical nature of that project prevented him, as a philosopher, from intervening in the species-ethical discourse opened up by these issues. Postmetaphysical philosophy cannot supply visions of the good, or make claims about the correct form of human life. His paradigm was at the mercy of discourses from which it had withdrawn. Habermas could have ignored the anthropic problem, or he could have abandoned

---

[1] See Euripides, "Alcestis*"*, in Paul Roche ed., *Three Plays of Euripides* (New York: Norton, 1974).

postmetaphysical thinking and addressed the species-ethical discourse directly. Instead, he chose to bring a god onstage.

By adopting his postsecular stance, Habermas maintains his commitment to postmetaphysical thinking, and at the same time gains access – in "translation" – to the ethical and metaphysical concepts which he needs in order to combat hard naturalism and genetic modification. The timing of the postsecular turn – which as I have shown, took place after Habermas had exhausted every attempt to make a postmetaphysical case against genetic modification – corroborates this interpretation. So does the fact that every single example of sacred-to-secular translation in Habermas' writings involves an argument against stem-cell research – translating "God created man in his own image" into "the gamete fertilized ex utreo is a subject of human rights." (FHN 109, 114-5)

Thanks to Heracles' intervention, Alcestis gets to make her tragic sacrifice, and to live happily ever after. The manoeuvre is ad hoc and inelegant, but it rescues the play.

Residual Atheism

The origins of the postsecular turn, then, lie in the postmetaphysical nature of Habermas' paradigm. He aims to do philosophy without making any metaphysical claims. This would be a reasonable programmatic choice for a philosopher with a limited project, or a narrow definition of metaphysics. Habermas has neither. His project is broad in the extreme, touching on topics such as meaning in language, metaethics, political legitimacy, the nature of reason, and the evolution of society. His definition of "metaphysics" is also broad: it includes questions of the good and philosophical anthropology, as well as ontology. In practice, he cannot avoid making some assumptions which are, in his terms, metaphysical.

Habermas' paradigm is based on discourse. It is from the discursive interactions of human beings that he rationally reconstructs his theories of meaning and validity in language, of morality, and of democracy. All of this relies on the unstated assumption that humans are, and should be, rational language-users of a certain type. When factors beyond his control threatened to undermine this unstated account of human nature, Habermas was at a loss. His attempt to construct a philosophical discourse without "metaphysics" simply meant that, at the crucial moment, another, unashamedly metaphysical discourse took over. Ironically, Habermas' abstention from metaphysics simply prepared the ground for the return of religion.

Christopher Watkin discusses a related problem, encountered by approaches to philosophy which attempt to leave theological thinking behind. Much Enlightenment thought, he suggests, is characterised by *imitative atheism*. Imitative atheist thinkers reject the idea of God, but then simply replace it with a term such as "humanity" or "reason." These placeholders play the same role in their thinking as God did in theological thought, providing meaning, a ground for truth or morality, a source of transcendence, and so on. Imitative atheism is parasitic on religion. It retains its form, changing only its content.[2] *Residual* or *ascetic atheism* develops as a response to this parasitism. Residual atheists reject *any* way of thinking previously connected to religion – they engage in a "renunciation of any imitation of the dead God."[3] Rather than grounding truth and transcendence and the good in humanity or reason, residual atheism avoids these concepts altogether. The place previously occupied by God is now empty. The problem with residual atheism, as Watkin points out, is that it concedes too much to theological thinking, "confining itself to playing the role of religion's ascetic poor relation, prone to define itself in terms of what it lacks…"[4] As a result, it makes room for religion to return, and occupy

---

[2] Christopher Watkin, *Difficult Atheism* (Edinburgh: Edinburgh University Press, 2011), pp 1-3. Watkin gives Diderot, Comte and Feuerbach as examples of imitative atheist thinkers.

[3] Ibid, pp 3-4. Watkin gives Nietzsche, Camus (on Blanchot's reading), Foucault and, to an extent, Derrida as examples of residual atheistic thinkers.

[4] Ibid, p 11.

the ethical and metaphysical territory which it has abandoned. Residual atheism defines its role in bare, ascetic terms, and therefore cedes concepts such as transcendence, truth, and the good to theological thinking. When, as inevitably happens, people have need of these concepts, imitative atheism has nothing to offer them. They must turn to religion. Residual atheism, as Watkin puts it, "finds itself unable to refute the turn to religion; in fact it is the very seedbed in which the turn can take root and bloom."[5]

Postmetaphysical thinking, I suggest, suffers from the same problem as residual atheism.[6] It is precisely because Habermas wants to abstain from discourses about ethics, human nature and metaphysics that his paradigm ends up leaning on, and eventually yielding to, religion. This is already apparent in his procedure of ethical appropriation; it culminates in the postsecular turn. The fact that Habermas does not want to talk about metaphysical questions does not mean that these questions go away. A paradigm which begins by rejecting metaphysics, and ends by relying on it for support, is a failed paradigm.

So what can modern philosophy do, if abstaining from metaphysics is self-defeating? Watkin argues that truly post-theological thinking should neither imitate religion, nor withdraw from its territory. It should, in my terms, conceptually appropriate the contents of religious traditions, and make use of them as it sees fit. Watkin points to the philosophical use of traditionally religious terms such as "grace," "faith," "spirit" and "praise" by thinkers such as Alain Badiou, Jean-Luc Nancy and Quentin Melliassoux as examples. "What post-theological integration attempts," he says, "is not to oppose theism but to occupy it, not to expel theism

---

[5] Ibid, p 11.

[6] Of course, postmetaphysical thinking is a form of agnosticism rather than atheism. To assert the non-existence of God would be to make a metaphysical claim. Nonetheless, the structural parallels between Watkin's argument and my account of how the anthropic problem leads to the postsecular turn are clear.

but to ingest it, taking terms and patterns of thought previously associated with theism and re-inscribing them in a way which is not to be confused either with parasitism or with asceticism."[7]

For philosophers such as Habermas, and those of us influenced by him, the situation is a little different. The point is not to return to metaphysics, but to admit that we have been doing metaphysics all along. Rather than attempting to shrink philosophy down to a narrow band of immanent problem-solving, let us openly express the ethical and anthropological grounds of our philosophical projects. If "metaphysical" questions inevitably arise, not only in philosophy but in politics, then we must ask ourselves how we would like to see them answered. Should we, as philosophers, do our best to offer answers which are critical, rigorous and open to questioning? Or should we stand back, and let traditional sources of metaphysical authority, shielded from criticism by the aura of the sacred, take centre stage? Any philosophy worth doing will choose the former.

Pyrrhic Translation

Habermas, with his model of postsecular deliberative democracy, chooses the latter. In Part 3, I outlined some of the drawbacks of this model, chief amongst them the pyrrhic nature of sacred-to-secular translation. Previous commentators have cast doubt on the procedure on the basis that it is burdened with contradictory skopoi, which demand that its products must be recognisable to monolingually religious citizens, and at the same time secular enough for the formal public sphere. Using my findings in Parts 1 and 2 of this text, I go beyond these criticisms. My claim is that *no* version of the procedure – whether we interpret it as ethical appropriation, conceptual appropriation, or as analogous to interlinguistic translation – can play the role that Habermas wants it to play.

Suppose that we regard sacred-to-secular translation as a variety of appropriation, as Habermas himself seems to do. My accounts of the various types of appropriation in Part 2

---

[7] Ibid, p 14.

illustrate the problems with this approach. A type one ethical appropriation, which transposes an ethical value from metaphysical to postmetaphysical discourse along with its "substance of piety," would violate the skopos of maintaining the secular nature of the formal public sphere. A type two ethical appropriation, which embeds the ethical value in an ersatz substance of piety, would violate the skopos of acknowledgement. If we regard translation as conceptual appropriation, a similar problem occurs. A conceptual appropriation could not hope to meet the skopos of acknowledgement; in addition, it seems unreasonable to require citizens to engage in such a demanding procedure.

If we interpret sacred-to-secular translation as analogous to interlinguistic translation, then the problems afflicting it become insurmountable. Here I draw on Habermas' original account of sacred and profane language, which I outlined in Part 1. The crucial point is that sacred language relies on a "syndrome" of entangled validity claims, worlds and attitudes, while profane language is fully differentiated. This leads to two problems. Firstly, the structural differences between religious and secular language are such that attempts at translation will run into the problems of non-equivalent terms identified by Roman Jakobson. Since the source and target languages in this case are not as flexible as natural languages, Jakobson's solutions to these problems (such as coining neologisms and finding alternate sources of semantic information) cannot be used. Secondly, religious language, as Habermas understands it, displays form/content dependence. Religious concepts rely on the undifferentiated nature of sacred language. Attempting to put this content into another form simply sacrifices the content, as with wordplay or self-referential statements. Secular translations of sacred contents are invariably pyrrhic.

However we interpret it, sacred-to-secular translation is unworkable, and a model of deliberative democracy based on it is therefore unacceptable. It seems to me that Habermas' main aim in formulating it was to recruit religious concepts for the ethical-political struggle against genetic modification, rather than to ease the supposed cognitive burdens of religious

citizens. The awkward hybrid nature of the model – which, as Maeve Cooke puts it, marries a secular state to a postsecular society, with the dubious process of translation mediating between the two – attests to this. Habermas wants to make the smallest possible adjustment to secular deliberative democracy, consistent with gaining access to religious concepts. His insistence that the formal public sphere be purely secular, behind the hygienic cordon of the institutional translation proviso, indicates how limited and lukewarm his conversion to postsecularism is. Habermas still assumes that religions are below the threshold of rational discourse, fit only to be mined for their conceptual resources.

As a *deus ex machina* manoeuvre to solve the anthropic problem without abandoning postmetaphysical thinking, postsecular deliberative democracy makes sense. As a genuine attempt to, accommodate religious citizens, it does not. I do not know what model of democracy would be best at resolving the tension between monoglot public reason and the irreducible diversity of beliefs found in pluralistic societies; but whatever the answer is, we will not find it by translating the sacred.

Bibliography

Accetti, Carlo Invernizzi "Can Democracy Emancipate Itself from Political Theology? Habermas and Lefort on the Permanence of the Theologico-Political" *Constellations* 17:2 (2010)

Adams, Nicholas *Habermas and Theology* (Cambridge: Cambridge University Press, 2006)

Adorno, Theodor W. *Negative Dialectics* (New York and London: Continuum, 2007)

Aguirre, Javier "Habermas' account of the role of religion in the public sphere: A response to Cristina Lafont's critiques through an illustrative political debate about same-sex marriage," *Philosophy and Social Criticism* 39:7 (2013)

Aissen, Judith "On the Syntax of Obviation," *Language* 73:4 (1997)

Allen, Amy "Having One's Cake and Eating It Too: Habermas's Genealogy of Postsecular Reason," in Craig Calhoun, Eduardo Mendieta and Jonathan VanAntwerpen eds., *Habermas and Religion* (Cambridge: Polity, 2013)

Apel, Karl-Otto "Regarding the Relationship of Morality, Law and Democracy: On Habermas' *Philosophy of Law* (1992) from a transcendental-pragmatic point of view," in Mitchell Aboulafia, Myra Bookman and Catherine Kemp eds., *Habermas and Pragmatism* (London and New York: Routledge, 2002)

Arendt, Hannah *The Human Condition* (Chicago and London: University of Chicago Press, 1998)

Audard, Catherine "Rawls and Habermas on the Place of Religion in the Public Domain," in James Gordon Finlayson and Fabian Freyenhagen eds., *Habermas and Rawls: Disputing the Political* (New York and London: Routledge, 2011)

Audi, Robert "The Separation of Church and State and the Obligations of Citizenship," *Philosophy and Public Affairs* 18:3 (1989)

- The Place of Religious Argument in a Free and Democratic Society," *San Diego Law Review* 30:4 (1993)

Audi, Robert and Wolterstorff, Nicholas *Religion in the Public Square* (Lanham, Md.: Rowman and Littlefield, 1997)

Austin, John L. *How To Do Things With Words* (Oxford: Clarendon, 1962)

Barbato, Mariano and Kratochwil, Friedrich "Towards a Post-Secular Political Order?" *European Political Science Review* 1:3 (2009)

Baumeister, Andrea "The Use of Public Reason by Religious and Secular Citizens: Limitations of Habermas' Conception of Religion in the Public Realm," *Constellations* 18:2 (2011)

Baxter, Hugh *Habermas: The Discourse Theory of Law and Democracy (Jurists: Profiles in Legal Theory)* (Stanford: Stanford University Press, 2011)

Benjamin, Walter "The Work of Art in the Age of Mechanical Reproduction," and "Theses on the Philosophy of History," in Hannah Arendt ed., *Illuminations* (London: Pimlico, 1999)

Bernstein, Richard J. "Naturalism, Secularism and Religion: Habermas' Via Media," *Constellations* 17:1 (2010)

Bertland, Alexander "Habermas and Vico on Mythical Thought," in Lewis Edwin Hahn ed., *Perspectives on Habermas* (Chicago: Open Court, 2000)

Boettcher, James W. "Habermas, Religion and the Ethics of Citizenship," *Philosophy and Social Criticism* 35:1 (2009)

Boon, Vivienne "Jürgen Habermas and Islamic Fundamentalism: on the limits of discourse ethics," *Journal of Global Ethics* 6:2 (2010)

Brieskorn, Norbert, "On the Attempt to Recall a Relationship," in *An Awareness of What is Missing: Faith and Reason in a Post-Secular Age* (Cambridge: Polity, 2010)

Burge, Tyler "Self-Reference and Translation," in F. Guenther and M. Guenther-Reutter eds., *Meaning and Translation – Philosophical and Linguistic Approaches* (London: Gerald Duckworth, 1978)

Casanova, José *Public Religions in the Modern World* (Chicago: University of Chicago Press, 1994)

Chambers, Simone "How Religion Speaks to the Agnostic: Habermas on the Persistent Value of Religion" *Constellations* 14:2 (2007)

Clarke, Randolph "The Appearance of Freedom," *Philosophical Explorations* 10:1 (2007)

Cohen, Joshua "Moral Pluralism and Political Consensus," in David Copp, Jean Hampton, and John Roemer eds., *The Idea of Democracy* (Cambridge: Cambridge University Press, 1993)

Colclasure, David L. *Habermas and Literary Rationality* (London and New York: Routledge, 2010)

Cooke, Maeve *Language and Reason: A Study of Habermas' Pragmatics* (London/Cambridge, MA: MIT Press, 1997)

- *Re-Presenting the Good Society* (London and Cambridge, Mass.: MIT Press, 2006)
- "Salvaging and Secularizing the Semantic Contents of Religion: The Limitations of Habermas's Postmetaphysical Proposal" *International Journal for Philosophy of Religion* 60 (2006)
- "A Secular State for a Postsecular Society? Postmetaphysical Political Theory and the Place of Religion," *Constellations* 14:2 (2007)
- "Translating Truth," *Philosophy and Social Criticism* 37.4 (2011)
- "Violating Neutrality? Religious Validity Claims and Democratic Legitimacy," in Craig Calhoun, Eduardo Mendieta and Jonathan VanAntwerpen eds., *Habermas and Religion* (Cambridge: Polity, 2013)
- "The Limits of Learning: Habermas' Social Theory and Religion" *European Journal of Philosophy*, forthcoming.

Costanzo, W. V. "The French Version of 'Finnegans Wake': Translation, Adaptation, Recreation" *James Joyce Quarterly* 9:2 (1972)

Dews, Peter *The Idea of Evil* (Oxford: Blackwell, 2008)

Durkheim, Emile *The Elementary Forms of Religious Life* (London: Allen & Unwin, 1976)

Duvenage, Pieter *Habermas and Aesthetics: The Limits of Communicative Reason* (Cambridge: Polity, 2003)

Eder, Klaus *The Social Construction of Nature* (London: Sage, 1996)

Edgar, Andrew *The Philosophy of Habermas* (Chesham: Acumen, 2005)

Ennis, Phil "Habermas, Reason, and the Problem of Religion: The Role of Religion in the Public Sphere," *The Heythrop Journal* 48:6 (2007)

- "Habermas, Democracy and Religious Reasons," *The Heythrop Journal* 51:4 (2010)

Euripides, "Alcestis*",* in Paul Roche ed., *Three Plays of Euripides* (New York: Norton, 1974)

Finlayson, James Gordon, and Freyenhagen, Fabian "Introduction" to *Habermas and Rawls: Disputing the Political* (New York and London: Routledge, 2011)

Fisher, Andrew "Flirting with Fascism – the Sloterdijk Debate," *Radical Philosophy 99* (Jan/Feb 2000)

Frega, Roberto "Equal Accessibility to All: Habermas, Pragmatism, and the Place of Religious Beliefs in a Post-Secular Society," *Constellations* 19:2 (2012)

Goffman, Erving *The Presentation of Self in Everyday Life* (London: Penguin, 1990)

Graumann, Sigrid "Experts on philosophical reflection in public discourse - the German Sloterdijk debate as an example," *Biomedical Ethics* 5:1 (2000)

Habermas, Jürgen *Theory and Practice* (Boston: Beacon Press, 1974)

- "The Analytical Theory of Science and Dialectics" and "A Positivistically Bisected Rationalism," in Theodor W. Adorno ed., *The Positivist Dispute in German Sociology* (London: Heinemann, 1976)
- *Zur Rekonstruktion des Historischen Materialismus* (Frankfurt am Main: Suhrkamp, 1976)
- *Legitimation Crisis* (London: Heinemann, 1976)
- "History and Evolution" *Telos* 39 (1979)
- *Communication and the Evolution of Society* (London: Heinemann, 1979)
- *Knowledge and Human Interests* (London: Heinemann, 1981)
- "Questions and Counterquestions," in Richard J. Bernstein ed., *Habermas and Modernity* (Cambridge: Polity, 1985)
- *On the Logic of the Social Sciences* (Cambridge: Polity, 1990)
- *Moral Consciousness and Communicative Action* (Cambridge: Polity, 1990)
- *Justification and Application: Remarks on Discourse Ethics* (Cambridge: Polity, 1993)
- *Postmetaphysical Thinking: Philosophical Essays* (Cambridge: Polity, 1998)
- *The Inclusion of the Other: Studies in Political Theory* (Cambridge, Mass.: MIT Press, 2001)
- "A Conversation About God and the World: Interview with Eduardo Mendieta," in Eduardo Mendieta ed., *Religion and Rationality: Essays on Reason, God and Modernity* (Cambridge: Polity, 2002).
- *The Future of Human Nature* (Cambridge: Polity, 2003)
- *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy* (Cambridge: Polity Press, 2004)
- *The Theory of Communicative Action, Volume 1: Reason and the Rationalization of Society* (Cambridge: Polity, 2004)
- "Transcendence from Within, Transcendence in this World," in Eduardo Mendieta ed., *The Frankfurt School on Religion: Key Writings by the Major Figures* (London: Routledge, 2004)
- *Truth and Justification* (Cambridge, Mass.: MIT Press, 2005)
- *The Theory of Communicative Action, Volume 2: Lifeworld and System: A Critique of Functionalist Reason* (Cambridge: Polity, 2006)

- "The Language Game of Responsible Agency and the Problem of Free Will: How can epistemic dualism be reconciled with ontological monism?" *Philosophical Explorations* 10:1 (2007)
- *Between Naturalism and Religion: Philosophical Essays* (Cambridge: Polity, 2008)
- *Europe: The Faltering Project* Cambridge: Polity, 2009)
- *"A Reply," in *An Awareness of What is Missing: Faith and Reason in a Post-Secular Age* (Cambridge: Polity, 2010)
- "The Political - The Rational Sense of a Questionable Inheritance of Political Theology," in Eduardo Mendieta and Jonathan VanAntwerpen eds., *The Power of Religion in the Public Sphere* (New York and Chichester: Columbia University Press, 2011)
- "Reply to My Critics," in Craig Calhoun, Eduardo Mendieta and Jonathan VanAntwerpen eds., *Habermas and Religion* (Cambridge: Polity, 2013)

Harrington, Austin "Habermas and the 'Post-Secular Society,'" *European Journal of Social Theory* 10:4 (2007)

Harvey, Keith "Compensation," in Mona Baker ed., *The Routledge Encyclopaedia of Translation Studies* (London and New York: Routledge, 2001)

Heath, Joseph "What is a Validity Claim?" *Philosophy and Social Criticism* 24:23 (1998)

Hegel, G. W. F. *Lectures on the History of Philosophy* vol. III (New Jersey: Humanities Press, 1983)

- "Fragment on the "Life-Course" of God," in Jon Stewart ed., *Miscellaneous Writings of G. W. F. Hegel* (Evanston, Illinois: Northwestern University Press, 2002)
- *Outlines of the Philosophy of Right* (Oxford: Oxford University Press, 2008)

Heidegger, Martin *The Fundamental Concepts of Metaphysics: World, Finitude, and Solitude* (Bloomington: Indiana University Press, 1996)

Henrich, Dieter "What is metaphysics – What is modernity? Twelve theses against Jürgen Habermas," in Peter Dews ed., *Habermas: A Critical Reader* (Oxford/Malden, MA: Blackwell, 1999)

Honneth, Axel and Joas, Hans *Social Action and Human Nature* (Cambridge: Cambridge University Press, 1988)

How, Alan R. *The Habermas-Gadamer Debate and the Nature of the Social: Back to Bedrock* (Aldershot: Avebury, 1995)

- "Habermas, History, and Social Evolution: Moral Learning and the Trial of Louis XVI," *Sociology* 35:1 (2001)

Ingram, David *Habermas: Introduction and Analysis* (Ithaca and London: Cornell University Press, 2010)

Irlenborn, Bernd "Religion in the Public Sphere: Habermas on the Role of Christian Faith" *The Heythrop Journal* 53:3 (2012)

Jaspers, Karl *The Origin and Goal of History* (London: Routledge and Keegan Paul, 1953)

Jakobson, Roman "On Linguistic Aspects of Translation," in Rainer Schulte and John Biguenet eds., *Theories of Translation: An Anthology of Essays from Dryden to Derrida* (Chicago and London: University of Chicago Press, 1992)

Junker-Kenny, Maureen "Genetic Enhancement as Care or as Domination? The Ethics of Asymmetrical Relationships in the Upbringing of Children" *Journal of Philosophy of Education* 39:1 (2005)

- *Habermas and Theology* (London and New York: T & T Clark, 2011)

Kierkegaard, Søren *Either/Or: A Fragment of Life*, vols. 1 and 2 (London: Oxford University Press, 1946)

- *The Sickness Unto Death: A Christian Psychological Exposition for Upbuilding and Awakening* (Princeton, New Jersey: Princeton University Press, 1980)

Kirk, G. S. *The Nature of the Greek Myths* (London: Penguin, 1974)

Knuth, Leo "The *Finnegans Wake* Translation Panel at Trieste" *James Joyce Quarterly* 9:2 (1972)

Kohlberg, Lawrence "From Is To Ought: How to Commit the Naturalistic Fallacy and Get Away with It in the Study of Moral Development," in Theodore Mischel ed., *Cognitive Development and Epistemology* (New York/London: Academic Press, 1971)

Koller, Werner "Equivalence in Translation Theory," in Andrew Chesterman ed., *Readings in Translation Theory* (Helsinki: Oy Finn Lectura Ab, 1989)

Laden, Anthony Simon "The House That Jack Built: Thirty Years of Reading Rawls," *Ethics* 113:2 (2003)

Lafont, Cristina "Religion in the Public Sphere: Remarks on Habermas's Conception of Public Deliberation in Post-Secular Societies" *Constellations* 14:2 (2007)

- "Religion and the Public Sphere: What are the Deliberative Obligations of Democratic Citizenship?" *Philosophy and Social Criticism* 35:1 (2009)

Laitinen, Lea "Zero Person in Finnish: A grammatical resource for construing human reference," in Marja-Lissa Helasvuo and Lyle Campbell ed., *Grammar from the Human Perspective: Case, Space and Person in Finnish* (Amsterdam: John Benjamins, 2006)

Larmore, Charles "The Foundations of Modern Democracy: Reflections on Jürgen Habermas," *European Journal of Philosophy* 3:1 (1995)

Levi, Primo *The Drowned and the Saved* (London: Abacus, 2003)

Lima, María Herrera "The Anxiety of Contingency: Religion in a Secular Age," in Craig Calhoun, Eduardo Mendieta and Jonathan VanAntwerpen eds., *Habermas and Religion* (Cambridge: Polity, 2013)

Low, Peter Alan "Translating Jokes and Puns," *Perspectives* 19:1 (2011)

Luhmann, Niklas *The Differentiation of Society* (New York: Columbia University Press, 1982)

Magee, G. A. *Hegel and the Hermetic Tradition* (Ithaca and London: Cornell University Press, 2001)

Marx, Karl "On the Jewish Question," in T. B. Bottmore ed., *Karl Marx: Early Writings* (London: C. A. Watts, 1963)

- *A Contribution to the Critique of Political Economy* – Preface," in *Later Political Writings*, ed. Terrell Carver (Cambridge: Cambridge University Press, 1996)

McCarthy, Thomas "Reflections on Rationalization in *The Theory of Communicative Action,*" in Richard J. Bernstein ed., *Habermas and Modernity* (Cambridge, MA: MIT Press, 1988)

- *Ideals and Illusions* (Cambridge, Mass.: MIT Press, 1991)
- "The Burdens of Modernized Faith and Postmetaphysical Reason in Habermas's 'Unfinished Project of Enlightenment'," in Craig Calhoun, Eduardo Mendieta and Jonathan VanAntwerpen eds., *Habermas and Religion* (Cambridge: Polity, 2013)

Mead, George Herbert *Mind, Self and Society* (Chicago: Chicago University Press, 1967)

Mendieta, Eduardo "Religion in Habermas' Work," in Craig Calhoun, Eduardo Mendieta and Jonathan VanAntwerpen eds., *Habermas and Religion* (Cambridge: Polity, 2013)

Mulhall, Stephen *Philosophical Myths of the Fall* (Princeton, NJ: Princeton University Press, 2005)

Nelson, Benjamin *On the Roads to Modernity: Conscience, Science, and Civilizations – Selected Writings* ed. Toby E. Huff (Totowa, New Jersey: Rowman and Littlefield, 1981)

Nietzsche, Friedrich *The Twilight of the Idols and The Antichrist* (Harmondsworth: Penguin, 1990)

- *Thus Spoke Zarathustra* (Oxford: Oxford University Press, 2005)

Oakley, Francis *Kingship* (Oxford: Blackwell, 2006)

Owen, David "Habermas' Developmental Logical Thesis: Universal or Eurocentric?" *Philosophy Today* 42:1 (1998)

Patberg, Markus "Supranational Constitutional Politics and the Method of Rational Reconstruction," *Philosophy and Social Criticism* 40:6 (2014)

Pedersen, Jørgen "Habermas' Method: Rational Reconstruction," *Philosophy of the Social Sciences* 38:4 (2008)

Pensky, Max "Solidarity with the Past and the Work of Translation: Reflections on Memory Politics and the Postsecular," in Craig Calhoun, Eduardo Mendieta and Jonathan VanAntwerpen eds., *Habermas and Religion* (Cambridge: Polity, 2013)

Piaget, Jean *Play, Dreams and Imitation in Childhood* (London: Routledge & Keegan Paul, 1962)

Picard, Barbara Leonie *Norse Gods and Heroes* (London: Oxford University Press, 1961)

Plato, *Statesman* (Cambridge: Cambridge University Press, 1995)

Pound, Ezra "Guido's Relations," in Rainer Schulte and John Biguenet eds., *Theories of Translation: An Anthology of Essays from Dryden to Derrida* (Chicago and London: University of Chicago Press, 1992)

Prusak, Bernard G. "Rethinking 'Liberal Eugenics': Reflections and Questions on Habermas on Bioethics." *Hastings Center Report* 35:6 (November-December 2005)

Puntel, Lorenz B. "Habermas' Postmetaphysical Thinking: A Critique," (www.philosophie.uni-muenchen.de)

Rawls, John "The Priority of Right and Ideas of the Good," *Philosophy and Public Affairs* 17:4 (1988)

- *Political Liberalism* (New York: Columbia University Press, 1996)
- "The Idea of Public Reason Revisited," in *The Law of Peoples* (Cambridge, Mass. and London: Harvard University Press, 2001)

Rehg, William *Insight and Solidarity: The Discourse Ethics of Jürgen Habermas* (Berkley: University of California Press, 1997)

Ricoeur, Paul "On the Paradigm of Translation" and "A 'Passage': Translating the Untranslatable," in Eileen Brennan ed., *On Translation* (London and New York: Routledge, 2006)

Ricken, Friedo, "Postmetaphysical Reason and Religion," in *An Awareness of What is Missing: Faith and Reason in a Post-Secular Age* (Cambridge: Polity, 2010)

Rorty, Mary Varney "For Love of the Game: Peter Sloterdijk and Public Controversies in Bioethics" (http://web.stanford.edu/~mvr2j/for_love.html)

Ryle, Gilbert "Knowing How and Knowing That: The Presidential Address," *Proceedings of the Aristotelian Society*, New Series, vol. 46 (1945)

Schäffner, Christina "*Skopos* Theory," in Mona Baker ed., *The Routledge Encyclopaedia of Translation Studies* (London and New York: Routledge, 2001)

Schleiermacher, Friedrich "On the Differing Methods of Translation," in Rainer Schulte and John Biguenet eds., *Theories of Translation: An Anthology of Essays from Dryden to Derrida* (Chicago and London: University of Chicago Press, 1992)

Schmid, Michael "Habermas's Theory of Social Evolution," in John B. Thompson and David Held ed., *Habermas: Critical Debates* (London: Macmillan Press, 1982)

Searle, John *Speech Acts: An Essay in the Philosophy of Language* (Cambridge: Cambridge University Press, 1969)
- *Expression and Meaning: Studies in the Theory of Speech Acts* (Cambridge: Cambridge University Press, 1981)

Sellars, Wilfrid "Philosophy and the Scientific Image of Man," in *Science, Perception, and Reality* (Atascadero, CA: Ridgeview, 1991)

Sloterdijk, Peter *Sphären I – Blasen* (Suhrkamp, 1998)

- *Sphären II – Globen* (Suhrkamp, 1999)
- *Sphären III – Schäume* (Suhrkamp, 2004)
- "Rules for the Human Zoo: a response to the *Letter on Humanism*," *Environment and Planning D: Society and Space* 12 (2009)
- *Die nehmende Hand und die gebende Seite* (Suhrkamp, 2010)

Smith, Neil *Chomsky: Ideas and Ideals* (Cambridge: Cambridge University Press, 1999)

Strydom, Piet "The Ontogenetic Fallacy: The Immanent Critique of Habermas' Developmental Logical Theory of Evolution." *Theory, Culture and Society* 9 (1992)

- "Sociocultural Evolution or the Social Evolution of Practical Reason?: Eder's Critique of Habermas," *Praxis International* 13:3 (1993)

Taylor, Charles *A Secular Age* (Cambridge, Mass. and London: Belknapp, 2007)

Ungureanu, Camil "The Contested Relation Between Democracy and Religion – Towards a Dialogical Perspective?" *European Journal of Political Theory* 7:4 (2008)

Urbinati, Nadia "Laïcité in Reverse: Mono-Religious Democracies and the Issue of Religion in the Public Sphere" *Constellations* 17:1 (2010)

Walhof, Darren R. "Habermas, same-sex marriage and the problem of religion in public life" *Philosophy and Social Criticism* 39 (2013)

Watkin, Christopher *Difficult Atheism* (Edinburgh: Edinburgh University Press, 2011)

Weber, Max "The Social Psychology of the World Religions," in H. H. Gerth and C. Wright Mills ed., *From Max Weber: Essays in Sociology* (London: Routledge & Keegan Paul, 1970)

- *The Protestant Ethic and the Spirit of Capitalism* (London: Routledge, 2001)

Weithman, Paul J. *Religion and the Obligations of Citizenship* (Cambridge: Cambridge University Press, 2004)

Winch, Peter "The Idea of a Social Science" and "Understanding a Primitive Society," in Bryan Wilson ed., *Rationality (Key Concepts in Social Science)* (Oxford, Blackwell, 1970)

Wolterstorff, Nicholas "An Engagement with Jürgen Habermas on Postmetaphysical Philosophy, Religion, and Political Dialogue," in Craig Calhoun, Eduardo Mendieta and Jonathan VanAntwerpen eds., *Habermas and Religion* (Cambridge: Polity, 2013)

Yates, Melissa "Rawls and Habermas on Religion in the Public Sphere," *Philosophy and Social Criticism* 33:7 (2007)

- "Postmetaphysical Thinking," in Barbara Fultner ed., *Jürgen Habermas: Key Concepts* (Durham: Acumen, 2011)

Zaretsky, Eli "Hannah Arendt and the Meaning of the Public/Private Distinction," in Craig Calhoun and John McGowan (eds.) *Hannah Arendt and the Meaning of Politics* (London and Minneapolis: University of Minnesota Press, 1997)

The Bible. Authorized King James Version, ed. Robert P. Carroll and Stephen Prickett (Oxford: Oxford University Press, 2008)

"WIN-Gallup International, Global Index of Religiosity and Atheism 2012 – Press Release" (http://redcresearch.ie/wp-content/uploads/2012/08/RED-C-press-release-Religion-and-Atheism-25-7-12.pdf)