Science, Reason and Society: 
Foucault and Habermas

John McIntyre
Department of Philosophy
Faculty of Arts and Social Sciences
The University of Sydney
2017

A thesis submitted in partial fulfilment of requirements for the degree of Doctor of Philosophy
ABSTRACT

This thesis elucidates science as a social institution through the prism of a critical examination of the work of Michel Foucault and Jurgen Habermas. I consider the role of science in relation to all aspects of society: morality, religion and social institutions such as democracy and the economy. This requires a consideration of the operation of power, as well as processes of socialisation and the formation of particular forms of subjectivity. Science is viewed in relation both to all the other high cultural discourses such as philosophy and art as well as the everyday lifeworld and commonsense. Common commitments arise from the two thinkers’ relation to the tradition of the Enlightenment and its continuation in Critical Theory. Neither disputes the truth of modern science in general. They want to expose the false pretensions of various attitudes towards science and their power effects. They draw on broad contexts to view categories of thought which are revealed as not given by nature, but historically conditioned and distorted by power. Both recognise the need for a more reflexive perspective and see philosophy as able to articulate social problems not visible from the specialised perspectives of science.

Habermas wants to endorse cultural modernity by taking its knowledge and interpreting it for contemporary society, to show not only the limitations of science but, its emancipatory potential. By viewing science in harness with critical theory, he offers a developmental account, whereby the sciences are linked to cognitive advances of distanciation and differentiation. Foucault is more sceptical in his theorising, and more wary in his assessment of the human sciences and their ubiquitous power effects. His genealogical stance suspends commitment to science which he relativises as a “regime of truth”.

Rather than comparing and judging one or the other to be more correct, or seeking to reconcile their differences, I want to maintain the tension between the two projects such that we neither have to reconcile nor choose between them. If we attend not only to what Foucault and Habermas say about science, but how they employ science, we can see their projects as two different aspects of the self-reflexivity of modern thought, which both posits its own foundations and remains open to criticism.
This is to certify that to the best of my knowledge the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged. This thesis has not been submitted for any degree or other purpose.

John McIntyre 5/7/2017
Acknowledgements

I would like to express my profound gratitude to my supervisor Associate Professor John Grumley for his unwavering support and commitment to my research. John has shown great patience in carefully reading and commenting on numerous and voluminous drafts, offering suggestions and challenging my thinking in ways that has encouraged me to clarify my thoughts. I would also like to express my gratitude to Professor Paul Redding who, at several critical points was able to bring a fresh critical perspective. I also thank the staff of the Philosophy Department at Sydney University for the stimulating intellectual environment that their collective activities and interactions engender, particularly Associate Professor David MacArthur, whose enthusiasm for philosophy is infectious. Also thanks to my wife, Teresa Mok and my son Angus, for their support, patience and kindness.
**CONTENTS**

**PART I**

**INTRODUCTION**

<table>
<thead>
<tr>
<th>CHAPTER ONE</th>
<th>Modern science as a social institution - Foucault and Habermas</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>General Approach</td>
</tr>
<tr>
<td>2</td>
<td>Foucault’s Life, Motivations and Commitments.</td>
</tr>
<tr>
<td>3</td>
<td>Habermas’ Life, Motivations and Commitments.</td>
</tr>
<tr>
<td>4</td>
<td>Philosophy, Science and Modernity</td>
</tr>
<tr>
<td>5</td>
<td>Aim</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**PART II**

**FOUCAULT**

<table>
<thead>
<tr>
<th>CHAPTER TWO</th>
<th>Archaeology</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Science, History and Archaeology</td>
</tr>
<tr>
<td>2</td>
<td>Madness</td>
</tr>
<tr>
<td>3</td>
<td>The Birth of the Clinic</td>
</tr>
<tr>
<td>4</td>
<td>Order</td>
</tr>
<tr>
<td>5</td>
<td>Man and the Sciences</td>
</tr>
<tr>
<td>6</td>
<td>Concluding Remarks</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER THREE</th>
<th>Science and Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>From Archaeology to Genealogy</td>
</tr>
<tr>
<td>2</td>
<td>The Emergence and Growth of Power/Knowledge</td>
</tr>
<tr>
<td>4</td>
<td>The Constitution of Subjects, Concepts and Society</td>
</tr>
<tr>
<td>5</td>
<td>Bio-Power and Governmentality</td>
</tr>
<tr>
<td>6</td>
<td>Normative Confusions</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER FOUR</th>
<th>Ethics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>From Power to the Subject</td>
</tr>
<tr>
<td>2</td>
<td>Regimes of Truth</td>
</tr>
<tr>
<td>3</td>
<td>Ethics, Aesthetics and Spirituality</td>
</tr>
<tr>
<td>4</td>
<td>Genealogy of the Subject</td>
</tr>
<tr>
<td>5</td>
<td>Truth Subjectivity and Modernity</td>
</tr>
</tbody>
</table>
PART III
HABERMAS

CHAPTER FIVE       The Critique of Positivism to the Theory of Cognitive Interests     89
1  The Positivist Dispute                        89
2  Knowledge and Human Interests                 97
3  After the Theory of Cognitive Interests       106

CHAPTER SIX       The Theory of Communicative Action                      111
1  Formal Pragmatics                              113
2  Lifeworld and System                           115
3. Evolution and Rationalisation of Society       119
4  The Diagnosis of Modernity                      121
5  Fragmented Consciousness                        123
6  Colonisation of the Lifeworld                  126
7  Habermas’ Theoretical Structures               130

CHAPTER SEVEN  From the Theory of Communicative Action to the present        135
1  Between Facts and Norms.                      135
2  Truth and Justification                       139
3  Theory and Practice Revisited                 142
5  The Future of Human Nature                     145
5  Faith and Knowledge                            149
6  Free will and Determinism                      151
7  Science and Philosophy                         155

PART IV
CONCLUSION

CHAPTER EIGHT    Science Modernity and Freedom                       157
1  Self-referentiality and its modern radicalisation  157
2  Discovery and Self-Transformation            158
3  Participant and Observer                      162
4  Necessity and Contingency                      164
5. Norms and Foundations                         166
6. Power and Reason                               171
7.  Science, Religion and Philosophy             175

ABREVIATIONS                        181

BIBLOGRAPHY                        183
CHAPTER 1

Modern science as a social institution - Foucault and Habermas

A spectre is haunting modernity – the spectre of the question concerning science. Never before has the dream of consciously steering history seemed more compelling. And never before has so much hung on consciousness that is not fully conscious. Science will steer our future – neuroscience, genetic science, big data – but without more adequate reflection, will we steer science? As the dominant form of rationality in modern societies, science requires reflection. Granting a claim the honorific title of “science” usually obviates the need to consider “non-scientific” accounts. If such non-scientific accounts come into consideration, they still claim, or aspire, to not directly contradict the claims of science. To say that “science” is wrong, is to say that what we thought was science wasn’t really science, or certainly wasn’t good science, the very meaning of which is tied to truth. Like truth, “science” has a context-transcendent moment which allows it to surpass itself. Yet only science has the right to surpass science. Underlying this self-assured authority, there is a sense of disquiet, a suspicion that something is missing, particularly from the scientific account of man and society. It is not that science is incomplete, or that it may not be true. Rather, the concern arises from science’s relation to other discourses and practices and the possibility that its authority, if accepted without adequate reflection, could limit freedom and facilitate domination.

By pursuing its truth in terms of its own specialised conventions, science withdraws from the everyday experience of those whose understanding is filtered through an array of social discourses. In our modern rationalised culture, science appears to gives rise particular forms of rationality such as the means-ends rationality that Weber discussed (Weber 1978). Without wanting to be overdramatic, we could call this perception of remoteness and underhand influence a crisis of faith in the scientific worldview. By “crisis” I mean a situation which calls for an urgent response which we can’t provide, because we can’t adequately grasp this situation or the type of response required. In particular instances, the aims, methods or findings of science are clearly problematic in terms of unintended social consequences. For example, Thomas Szasz’s classic work criticised the propensity of psychiatry to reach into all aspects of life (Szasz 1974). More recently, Watters criticises the export of psychiatric diagnoses to societies where such “disorders” make no sense (Watters 2010).

One question that frequently lies behind such concerns is the role that social, economic and political power plays, not merely in the implementation of scientific findings, but the very framing of problems
as problems and the constitution of categories as objects of investigation. Given its power effects, can we insist on the value-freedom of science, even as an ideal, or does such insistence itself bolster a partisan position? Like questions about limits to scientific knowledge and forms of knowledge beyond science, such issues are far from settled. These questions are frequently revisited when new frontiers of science encroach on social life. The development of life sciences, such as genetic engineering and neuroscience, has consequences that seem to far outpace our considered understanding of them. The ensuing controversies focus not so much on the restricted truth claims of the sciences involved, but draw attention to what cannot be captured in scientific accounts of humans and their societies and the effects that the unconsidered endorsement and adoption of science has on these “objects”.

There is a temptation to think that it is only science, rather than religion, art, philosophy or common sense that offers reliable access to truth and secure knowledge. On this assumption, theories of secularisation and modernisation have simply assumed that religion will fade away under the clear light of an expanding scientific worldview. Problems to do with the interpretation and implementation of science are frequently partitioned off from “science itself” which is then seen as a “pure” neutral self-correcting project, oriented solely to truth as its ultimate value. Objections that science doesn’t give a fully adequate account of the world are then dismissed by the promise of the increasing rigor of the ongoing project of science to bring more accurate, detailed and comprehensive knowledge under its authoritative jurisdiction. This heroic optimism about science was seen paradigmatically in figures such as Popper, who viewed science, modelled on the natural sciences, as possessing an intrinsic potential for human liberation from natural and social constraints (Popper 1963). Whilst mid 20th century positivism is discredited by its reduction of all processes to physiological, physical or chemical events, it still lives on in different forms of reductive naturalism. This means that problems which inevitably have a social dimension are frequently addressed as merely technical problems.

Against such positive appraisals others, such as Horkheimer and Adorno, viewed the growth of scientific rationality as a narrowing of opportunities and a growing form of domination, a new mythology in which instrumental rationality comes to be seen as reason per se (Adorno & Horkheimer 1997). Later less radical thinkers, whilst not sharing this degree of pessimism, remain cautious or non-committal in their assessments of science. Others have taken the persistence of religious belief as reason to re-examine the Weberian conviction that the modern scientific worldview

---

1 A seminal secularisation theory is Max Weber’s *The Protestant Ethic and the Spirit of Capitalism* (Weber 2002).
2 For example the eliminative materialism in Churchland (1984)
3 For example, the naive enthusiasm for earth engineering as a response to climate change.
would inexorably lead to the decline of religion.\textsuperscript{4} The sort of history and philosophy of science initiated by Thomas Kuhn has rendered the “internal logic” of scientific progress deeply imbricated by “external influences”, leading to doubts about the neutrality and objectivity of science and, in more extreme forms, the truths of science.\textsuperscript{5} In recent decades, the sociology of scientific knowledge has sought the explanatory basis for scientific progress in the social and historical contexts of science.\textsuperscript{6}

These academic trends have been accompanied by a more general scepticism about, and even a fear of, and resistance to, science which seems a reaction against the faith invested in science by earlier generations. On one hand, vested interests which benefit from the continuation of certain modes of industrial production often see science as a threat.\textsuperscript{7} Yet these interests often respond cynically by citing other specialised areas of science or resorting to pseudo-science. Thus we can see economic arguments pitted against environmental arguments, both based on their respective sciences. On the other hand, vague misgivings about scientific rationality are often backed by the resistance of counter-movements that seek to defend ways of life against what appears as the inexorable force of instrumental modes of thinking.\textsuperscript{8} Yet these counter-movements also frequently draw on the very terms of the rationality they resist, suggesting a limit to what is thinkable and sayable beyond the form of rationality whose pre-eminent expression is science.

Michel Foucault (1926 – 1984) and Jurgen Habermas (1929 - ) illuminate these misgivings without recourse to simplistic judgements. Both are concerned with science as a historical phenomenon within the larger picture of knowledge and social life. Their nuanced and differentiated analyses elucidate the dangers and risks that science poses, whilst not denying its emancipatory possibilities. Whilst committed to scientific standards of evidence and objectivity, they challenge the uncritical acceptance of science as the paradigm of modern knowledge. Yet there are differences. Foucault wants to radically challenge Enlightenment humanism, whilst Habermas wants to recover emancipatory aspects latent within it. Both articulate their projects in their own terms, Habermas in terms of the public sphere, democracy and the ideals of communicative action and Foucault in terms of his genealogical histories of the human sciences.

\textsuperscript{4} As I will discuss in Chapter 7, Habermas was prompted by the post 9/11 political scene to reconsider the place of religion in “post-secular society”. see FHN pp. 101

\textsuperscript{5} Kuhn’s \textit{The Structure of Scientific Revolutions} created considerable controversy in the English speaking world, leading to extravagant misinterpretations of Kuhn’s position. Within the continental tradition, thinkers such as Foucault’s mentors, Bachelard and Canguilhem, provided a more nuanced and developed position several decades earlier (Kuhn 1996).

\textsuperscript{6} For example, Barnes, Bloor & Henry (1996).

\textsuperscript{7} For example, denial of anthropogenic climate change benefits certain industries by delaying actions which are perceived to devalue assets in those industries.

\textsuperscript{8} Such resistance is seen, for example, in campaigns against globalisation.
To date there have been journal articles on Foucault and science (Garland 1992; Garland 1997; Power 2011; Borch 2015; Voruz 2011; Bracken & Thomas 2010) and two books on Foucault’s early work in relation to science (Webb 2013; Gutting 1989). There have also been journal articles on Habermas and science and a number of books which tend to consider how his work bears on specific aspects and areas of science (Alföldi 1985; Rehg 2009). There is clearly a gap in the literature. This gap is significant because Habermas and Foucault are two towering figures of late 20th century continental philosophy, who in different but overlapping ways, analyse science from perspectives which challenge conventional understandings of its relation to contemporary social reality. A work which analyses, compares and critically evaluates what they have to say about the dominant form of contemporary knowledge is indispensible.

Following the so-called “debate” between Foucault and Habermas, commentators tended to look for reasons to support or reject their respective positions. However, in recent decades, Koopman, Cooke, Allen and others have sought to accommodate their best insights (Allen 2008; Cooke 2006; Koopman 2013). I don’t intend to either adjudicate or reconcile, but simply account for the tension between their two projects in a way that doesn’t require us to choose between them, or smooth over the tensions. This involves drawing on these two philosophers to elucidate the relations between science as an institution and form of knowledge and its environing social and historical context. I will consider the complex ways in which science and its context interact to condition, express and transform each other. This will include, for example, a consideration of the role and function of science in relation to aspects of culture such as law and morality and social institutions such as the family or religion. The analysis of interactions will require a consideration of the operation of power both generally and in specific forms such as economic and administrative power, as well as processes of socialisation and the formation of particular forms of subjectivity. The relation of science to other discourses such as philosophy, religion and everyday commonsense will be considered, with particular consideration given to the role of philosophy. I will tease out the conjecture of a broader framework, which reveals the positions of Foucault and Habermas as two distinct tendencies within the structure of modern self-reflexive thought. These positions are revealed not only by what Foucault and Habermas say directly about science, but how they use it and the role they grant science in their own critiques of reason.

We will see Habermas employ science to build a non-metaphysical foundation for a progressive account of reason, yielding an explanatorily powerful, yet fallible diagnosis of modernity as well as conceptual tools of analysis and prescription. We will see Foucault suspend science’s authority, adopt agnosticism towards its truth, so bringing it into view as one form of knowledge amongst others. The two strategies are not opposed. We need to both posit context-transcending ideals and reveal their illusory status as emerging from power-laden contexts. For both thinkers this is a political project in
which they seek to change consciousness by revealing it to be conditioned by a vast range of historical and social factors.

In this Chapter, I will set the scene for this investigation by firstly charting my approach in terms of how I will employ the two thinkers to bear on my topic, given their significant differences and commonalities. I will then discuss firstly Foucault’s, then Habermas’, motivations and commitments, drawing on the contexts, circumstances and events which shaped their lives - the politics, intellectual milieu, reading, teachers which shaped their projects. I will then turn to the imbricated relationship of philosophy to science, and how the two thinkers see this relationship. I will conclude with an outline of chapters.

1 GENERAL APPROACH

The informal debate between Foucault and Habermas may suggest that they were diametrically opposed on a range of matters.9 A considerable body of current literature now continues that debate. Foucault is typically criticised for various forms of relativism and eliminating the subject, Habermas is typically criticised for idealisations, abstract universalism and naturalism. Foucault and Habermas have frequently been seen as two opposed leftist philosophers, one humanist, the other non-humanist, both actively engaged with politics, though with differing assumptions and objectives. This impression was encouraged by the intermittent criticisms that they offered of each other’s work, which sometimes simply misunderstood what the other was claiming.

What is not always recognised is the extent to which Foucault and Habermas share common ground. For both Foucault and Habermas what is distinctive about philosophy is its anarchic nature, unconstrained by the methods of science which, in what Kuhn calls “normal science” appear relatively fixed and laid down in advance.10 They both draw on broad historical contextualisations to open up a critical perspective to view contemporary scientific categories of thought as not given by nature, but as historically conditioned, and therefore subject to change. Both expose the false pretensions of various versions of scientism and their oppressive and alienating power effects. Both recognised the need for a more reflexive perspective and see a form of critical philosophy as able to articulate social problems not visible from the specialised perspectives of science which now must be seen to have a more limited, though still important, scope.

---

9 The documents of the Habermas Foucault debate are mostly collected in Kelly (1992).
10 By “normal science” Kuhn is referring to the periods in which the paradigms which underlay basic assumptions are generally agreed upon and science works to solve problems within these accepted frameworks (Kuhn 1996)
In particular, Foucault and Habermas both see themselves as within the thread of the Enlightenment tradition culminating in Critical Theory which aims to integrate the social sciences into a broader context of reflective analysis and critique of society and culture. They aim to change society by understanding it, freeing human beings from entrapment in systems of dependence or domination, both internal and external. Habermas is a second generation critical theorist, taking on additional resources such as American pragmatism and analytic philosophy of language to guide work on intersubjective communicative rationality. Late in his career, Foucault recognised Critical Theory as a form of reflection close to his own. However, he doesn’t see his project as lining up exactly with the Frankfurt School, which he characterised (incorrectly) as wanting to realise an authentic human nature unencumbered by power. Whilst Foucault rejects the idea of an unconditioned subject, he still thinks we can exercise freedom by bringing to consciousness the historically contingent nature of constraints which we simply accept as given and natural. Like Habermas, he thinks that philosophy must understand its position within the particular social and historical formation of modernity, which it attempts to understand, diagnose and change. This diagnosis of the present involves, amongst other things, a critique of modernity’s paradigmatic form of reason, science.

Neither Habermas nor Foucault dispute (and I certainly don’t want to dispute) the truth claims of science, or rather what science narrowly defines as truth in terms of its particular rationality. The problem is that within modernity, scientific rationality has come to be seen as rationality per se. To gain a perspective on this problem, one needs to step outside the social and historical position of current science’s “truth regime” (to use Foucault’s term) to see science from a distance, from the outside as it were. Both Foucault and Habermas adopt such distanced standpoints from which to view contemporary society and its relations to science. Foucault employs genealogy to reveal the contingency of the categories of thought which we feel necessarily reflect the way the world really is. Habermas incorporates a scientific approach - a reconstructed defeasible empirical theory of social evolution – to provide a normative standpoint.

---

11 Critical theory was first defined in these terms in Max Horkheimer’s 1931 inaugural address (Horkheimer 1993)
12 In a 1978 Interview Foucault, referring to the members of the Frankfurt School, says he “should have read them long before, should have understood them much earlier” (EW3, 274). Foucault locates himself in the tradition of the critique of reason going through Kant, Nietzsche, Weber and Horkheimer. Foucault (1997, 97-120)
13 Foucault criticises the Frankfurt School’s conception of the subject as “permeated by Marxist humanism” and “connected to certain Freudian concepts, such as the relation between alienation and repression, between liberation and an end to alienation and exploitation …” (EW3, 274-5).
14 Drawing inspiration from Nietzsche’s Genealogy of Morals, Foucault’s genealogies question what the human sciences have simply assumed by connecting discourses to mechanisms of power, explaining the origin of what we take to be natural and given in terms of petty, contingent forgotten events. Genealogy thus undermines certainty by an account of historical beginnings “capable of undoing every infatuation” (NGH p.372). It does not seek to refute concepts as fictitious, but rather strips them of false authority.
My project can most broadly be described as a meta-critique – a critique of Habermas’ and Foucault’s critiques of science. As such it will ultimately bear on science itself as a social institution. Rather than accepting current science and its categories of analysis in terms of an internal logic oriented to truth or empirical adequacy, I will tease out Foucault’s and Habermas’ analyses in their terms – i.e. science as part of a history of social institutions, discourses and concepts. This means, for example, not analysing society’s ills from a level which accepts institutions such as asylums, schools and prisons and the current scientific discourses which justify them as simply given. Foucault and Habermas view such institutions, discourses and concepts as products of more profound historical transformations which themselves require analysis. To this effect Habermas will develop a dual model of “lifeworld” and “system” along with a historical analysis of their interactions which bring into view various social pathologies.¹⁵ Foucault will initially analyse discourses in terms of epistemes, the unconscious structures of particular epochs which underlie the production of scientific knowledge. Later, he will analyse knowledge as inextricably linked to power, found in the finest interstices of relations between subjects.¹⁶

With my emphasis on the human sciences, the question of the nature of man is clearly central to this work.¹⁷ This is the question that Foucault places firmly on the table in *The Order of Things*. Throughout all phases of his work, the notion of an essential nature of man, as an object of science, remains under challenge. It also comes under challenge in Habermas’ analyses of neuroscience and genetic manipulation which argues that fundamental distinctions and categories of thought are necessarily presupposed and can’t be eliminated. A related question will be the possibility of a science of society that, having addressed the essential nature of man, discerning his needs and legitimate aspirations, can lay down in advance what is required for the organisation of the good society.

Foucault and Habermas reject such sciences. Rather than a science of society, they offer a critical diagnosis. In this diagnosis, philosophical enquiry must work in tandem with the sciences.

To grasp how their diagnoses bears on science as a social institution, it will be essential to make sense of the work of the two thinkers as wholes. The difficulty is that different periods of work pull in different directions, and especially with Foucault, many commentators insist on three Foucaults. If

---

¹⁵ By “lifeworld”, Habermas means the implicit, though intuitively present, familiar web of presuppositions that have to be satisfied for an utterance to be meaningful or valid or invalid. By “system”, he is referring to society viewed in terms of various rule-bound structures, such as the market, which enable society to maintain itself by producing stable social patterns which transcend the intentions of individual actors.

¹⁶ See HS1, DP.

¹⁷ I have chosen not to adhere to a rigid distinctions between human, social, life or natural sciences and may occasionally stray into areas beyond the human sciences, as do both Foucault and Habermas.
early, middle and late Foucault are not in agreement which should we believe? Whilst this question is less pressing for Habermas, it is still important to grasp his work as a whole project.

With a body of publications, lectures and interviews extending over twenty-five years, it is hardly surprising to find major shifts in Foucault’s areas of concern, theoretical positions and approaches. These shifts have led to the conventional grouping of Foucault’s work into periods, amounting to a progression from the “archaeology” of the early period, through “genealogy”, to “ethics” of the late period. The archaeological period is generally considered to include *The Birth of the Clinic* (1963), *The Order of Things* (1966) and *The Archaeology of Knowledge* (1969). *The History of Madness* (1961) could be included as the beginnings of archaeology, although at a stage when Foucault was unable to formulate it clearly.18 Foucault’s genealogical works are generally considered to include *Discipline and Punish* (1975) and Volume 1 of *The History of Sexuality* (1976). His lectures from 1970 to 1978 at the *College de France* should also be included within his genealogical work. Foucault’s works on ethics include volumes 2 & 3 of *The History of Sexuality* (both 1984). The lectures from 1978 onwards should also be included in his ethics19

This periodisation can be situated within the broader context of the philosophical traditions and historical figures with which Foucault engages. Three major historical influences on Foucault were Kant, Nietzsche and the Cynics. Foucault continually engaged with the work of Kant, beginning with his supplementary doctoral thesis, a translation and commentary on Kant’s *Anthropology*. Later, in his archaeological period, *The Order of Things* analyses the legacy of Kant’s transcendental philosophy. Here Foucault’s recasts Kant’s critical project, by refusing the transcendental turn and insisting on the historicity of the conditions of possibility of knowledge or subjectivity. Still later, he wrote on Kant’s minor works including Kant’s essay *What is Enlightenment?* in which he claimed Kant as the key figure in the birth of the modern era and identified aspects of Kant’s work as part of a stream of philosophy that critically engages with the contemporary world, a stream in which Foucault saw himself.

---

18 By “archaeology” I mean an historical inquiry into the conditions of possibility for language or thought. These conditions, which Foucault calls the “historical *a priori*”, are relatively fixed within any historical epoch. Archaeology aims to describe discourses in terms of the conditions of their emergence and transformations by analysing statements, their actual occurrences and their effects. However, it eschews the quest for hidden meanings as expressions of subjects.

19 Foucault’s ethics refers to the ethical formation of the self by the self. Human subjectivity isn’t just formed within the constraints of the historical *a priori* of archaeology, or the power relations of his genealogies. Subjects also work on themselves by “technologies of the self.” “Ethics” for Foucault refers to the manner in which subjects form themselves as moral subjects.
Whilst Foucault described *The History of Madness* as composed “under the sun of the great Nietzschean quest”, Nietzsche was foregrounded only later in Foucault’s genealogies (HM, xxx). Foucault was interested in Nietzsche as a genealogist, who problematised truth as intimately entwined with relations of power, who understood subjectivity as a construct and saw a complexity of forces as the lowly origin of our concepts. In many respects Foucault went beyond Nietzsche, with a more sophisticated analysis, with greater archival breadth and attention to actual practices. “The only valid tribute to thought such as Nietzsche’s is precisely to use it, to deform it, to make it groan and protest” (PK, 53-4). I will also argue Foucault is less normatively ambitious.

In the last period of work, ancient philosophy, especially the Cynics, opened up fresh insights. After *The History of Sexuality* volume 1 Foucault began to “reorganise the whole study around the slow formation, in antiquity of a hermeneutics of the self” (HS2, 6). Although not trained as a classicist, but with the benefit of interaction with figures such as Pierre Hadot and Paul Veyne, Foucault conducted a series of courses and completed two more volumes of *The History of Sexuality*, all dealing with classical and post-classical antiquity. In this work, Foucault uncovered a new mode of subjectification, subjects working on themselves as ethical agents. He identified the Cynics’ radicalisation of philosophical practice as a way of life as an important event in that stream of philosophy, marginalised in modernity, that he came to recognise as his own.

Perhaps the deepest yet most indeterminate influence throughout Foucault’s entire work was Heidegger, only mentioned in his final interview, weeks before his death: “For me Heidegger has always been the essential philosopher... My entire philosophical development was determined by my reading of Heidegger. … I have never written anything on Heidegger … I think it is important to have a small number of authors with whom one thinks, with whom one works, but about whom one does not write” (Foucault 1989a, 470). It is not surprising that Foucault would read Heidegger whose *Being and Time* was the source of the work of Sartre, the major intellectual of France’s post-war years. It is difficult to locate the exact nature of Heidegger’s influence, since Foucault’s other references to Heidegger amount to little. However, one could surmise that Heidegger’s notion of the epochal nature of truth, which “holds complete dominion over all the phenomena that distinguish the age” influenced Foucault’s notion of *episteme* by which a shared unconscious framework conditions what can be said and thought within an epoch (Heidegger 1977, 115). In his genealogies, Foucault thematises power in this role of conditioning the possibility of both discursive and non-discursive practices. Whilst Heidegger emphasized the tendency of technology towards total ordering, Foucault refers to the totalizing tendency of power by "normalization." In modern societies, norms are progressively brought to bear on all aspects of life. This modern power is something entirely new - diffuse, continuous, invisible and constantly colonizing new domains. Both Foucault and Heidegger thought that individuals in modern society are determined by pervasive technological structures which
objectify and order social life. Both thought that by analysing the history of these unconscious structures we could come to recognise the contingency of our own thinking and transform ourselves by thinking differently.

The task in coming to terms with Habermas doesn’t, as with Foucault, involve making sense of the discontinuities suggested by periodisation. Rather it requires picking out the developmental strands and major commitments that run continuously through his entire project. Habermas certainly made changes, for example from an interest-based quasi-transcendental epistemology to a theory of communicative action as a basis of rationality. He also revised his consensual notion of truth to a notion involving reference to an external world. With well over 100 books published, Habermas’ project could be broken down into a number of overlapping research programs. Finlayson for example, suggests the following: a theory of communicative rationality, social theory, discourse ethics, democratic and legal theory and political theory (Finlayson 2005). Other ways to categorise Habermas vast output are plausible. If we look for guidance to a tradition, we can see that Habermas, as part of the evolving critical theory tradition was deeply connected to German idealism. But he also engaged with Anglo-Saxon analytical philosophy, speech act theory, pragmatism and American and European sociology. He drew on the major figures from American and European sociology, economics, psychoanalysis, hermeneutics.

As an actively engaged public intellectual Habermas spoke and wrote about a vast number of contemporary issues including politics, environment, genetic cloning, religion, the German past, the future of the nation-state, unification of Germany, constitutional law and the EU. He supported the students’ democracy movement in 1960s which opposed the German conservatives, the American led cold war and the Vietnam war, but ceased support when he felt the movement was losing its way. He shifted from his early criticism of positivism, technocracy and systems theory, as these targets became irrelevant, to criticise a new generation of allegedly conservative romantics – the postmodernists. With such a vast output of work covering so many areas, crossing so many disciplinary boundaries and drawing on so many traditions, it will be necessary to discern the motivating commitments and relevance to science that permeate the development of his project.

My task will be to make sense of both Foucault’s seeming discontinuity and Habermas’ diverse profusion in terms of overall projects which involve critical reflections on the role of science within modern society and culture. This will, in my final chapter, lead me to place Foucault and Habermas in relation to each other as two aspects of modern self-reflective rationality. Since the life we live as particular individuals within social and intellectual milieus is the fertile ground for theoretical and intellectual exertions, I will now sketch the formative life experiences, philosophical influences,
traditions and historical figures which motivated Habermas’ and Foucault’s principal concerns and commitments.

2 FOUCALT’S LIFE, MOTIVATIONS AND COMMITMENTS.

Foucault was born in 1926, the son of a surgeon. He was academically gifted but troubled. Depressed and obsessed with thoughts of self-mutilation and suicide, his father sent him to a psychiatrist. Eribon cites a doctor who knew Foucault, suggesting that his condition arose from the distress coming to terms with his homosexuality, an unsurprising response given the homophobia of 1940’s France (Eribon 1991, 26). It is also not surprising that in addition to his studies in history and philosophy, Foucault developed an “obsessive interest” in psychiatry, psychology and psychoanalysis (Eribon 1991, 27). He went on to study and qualify in psychology and work in mental hospitals. His writings and political engagements continued to reflect a concern for marginal groups and the powers to which they were subjected. It is surely this experience that prompted his insights into the inextricable relationship of power to scientific knowledge. Whilst the meanings of power and knowledge differ, Foucault claims they cannot be analysed separately. And both are inextricably tied to the subject, which must also be drawn into this constellation. Thus Foucault refers to particular forms of subjectivity, such as homosexuality, madness or delinquency, as constituted by particular regimes of power-knowledge.

Foucault entered the Ecole Normale Superieure in 1946, a period extending to the late 1950’s in which existential phenomenology and Marxism dominated French intellectual life. His earliest works, Maladie Mentale et Personalite and his Introduction to Dream and Existence by Ludwig Binswanger were powerfully influenced, respectively, by Marxism and existentialist phenomenology. Rabinow and Dreyfus situate Foucault’s ongoing development as triangulated between phenomenology and the two reactions to it: structuralism and hermeneutics (Rabinow and Dreyfus 1983, xix-xxi).

Husserl had thought of phenomenology as a source of absolute certainty in its pure intuitions of essential meanings and thus an unshakable foundation for all knowledge, including science. As practiced by Husserl, phenomenology treats man as both totally an object and totally a subject. The transcendental ego gives meaning to all objects including its own body and its empirical personality, and the culture and history which it “constitutes” and which conditions its empirical self.

---

20 Miller notes “that rumours of homosexuality could and did break academic careers. Prejudice was backed by legislation” (Miller 1993, 30)

21 By phenomenology, I am referring to that body of work that issues from Edmund Husserl, is taken up in France by Jean-Paul Sartre and Maurice Merleau-Ponty, and focuses on the investigation of the meaning-giving activity of the transcendental ego.
Foucault rejects this conception of the meaning-giving transcendental subject outright: “If there is one approach that I do reject, however, it is that (one might call it, broadly speaking, the phenomenological approach) which gives absolute priority to the observing subject, which attributes a constituent role to an act, which places its own point of view at the origin of all historicity – which, in short, leads to a transcendental consciousness” (OT xv). Foucault defined himself against the existential phenomenology of Sartre, the master-thinker of his formative period. He rejected Sartre’s centralisation of the subject which he mocked as “transcendental narcissism” and criticised the role of the “universal intellectual”, widely attributed to Sartre after the Second World War, able to pronounce on society in terms of universal principles. However, Foucault didn’t reject all phenomenology. Heidegger’s phenomenology, which eschewed the subject-object division, remained a powerful influence.

Towards the end of the 1950’s, the limitations of the phenomenological priority given to the existential subject became increasingly obvious in terms of the problems of language and the unconscious. Both structuralism and hermeneutics can be seen as reactions against phenomenology which attempt to transcend the subject/object division. Structuralism seeks the most basic elements (such as concepts, actions, classes of words) and the hidden rules or laws by which they are combined.22 Structural analysis aims to uncover structures which are the objective and universal constituents of human thought, action or language. Hermeneutics also turns away from the phenomenological priority given to subjectivity as the locus or origin of meaning. Whilst retaining the analytic emphasis on meaning, hermeneutics locates meaning in socio-historical and cultural practices and texts, in other words, forms that are not reducible to a conception of the meaning-giving subject.23 It assumes deep or ultimate truths awaiting recovery by interpretation.

Foucault should be distinguished from these approaches. In contrast to phenomenology, he isn’t willing to assume a fully autonomous subject which ascribes meaning through its own activity. He is also adamant in distinguishing himself from structuralism. Its purported objectivity and universalism is precisely what Foucault wants to avoid. In the foreword to the English edition of *The Order of Things*, he complains of “certain half-witted ‘commentators’ [who] persist in labelling me a ‘structuralist’. I … have used none of the methods, concepts, or key terms that characterise structural

---

22 By structuralism I am referring to a broad approach, which aspires to scientifcacy and is employed across disciplines as diverse as linguistics, anthropology and psychology. Common to these varied applications is a conviction that surface events and phenomena are to be explained by structures, data, and phenomena below the surface, which the structuralist method uncovers. The explicit and obvious is to be explained by and is determined by what is implicit and not obvious.

23 Hermeneutics also sets itself apart from structuralism, seeing the objective world described and analysed by structuralists as a product of human consciousness and its interpretive processes. Rather than structuralism’s notion of fixed and objective structures, hermeneutics allows for differences in interpretations and local, and often highly specific, readings of texts.
analysis” (OT, xv). Whilst retaining the displacement and decentering of meaning and subjectivity found in structuralism, he favoured historical contingency, rejecting structuralism’s construction of a formal rule-governed model of human behaviour. Whilst Foucault employs hermeneutic approaches, he does not regard this as granting access to ultimate truths. In all approaches Foucault would reject anything that put itself forward as a foundation for human knowledge, whether that be Husserl’s pure intuitions of essential meanings, structuralism’s universal laws or hermeneutics’ deep meanings.

But whilst Foucault rejects all these approaches, and contrasts are indeed strong, there are also points of engagement, certain similarities and points at which Foucault acknowledges an influence. For example he acknowledges that in writing The History of Madness he “accorded far too great a place, and a very enigmatic one too, to what I called, an ‘experience’, thus showing to what extent one was still close to admitting an anonymous and general subject of history” (AK, 18).\(^{24}\) Just as The History of Madness came close to phenomenology, the first printing of The Birth of the Clinic describes its approach as ‘structuralist’, and even though Foucault later retracted this, it does seem broadly structuralist in its analyses.

Whilst Rabinow’s and Dreyfus’ contextual ‘triangulation’ gives an adequate account of Foucault’s relationship to the dominant tendencies in French intellectual life, it fails to include what Foucault himself saw as his dominant influence. In particular, it fails to mention the history and philosophy of science, an independent sub-dominant stream in French intellectual life. On Foucault’s account, as structuralism came to be seen as superior to phenomenology, there was a series of attempts to combine structuralist or Freudian thought with Marxism, to produce a “structural-Freudo-Marxism” (Foucault 1989a, 350). Whilst this dominant approach continued until the late 1960’s, Foucault comments that there was a group of students in this period “who did not follow this movement [but instead]… participated in the history of science.” This group were aligned with the great historian of the life sciences George Canguilhem, “who had a decidedly influential effects on young French university life. Many of his students were neither Marxists, nor Freudians, nor Structuralists. … I’m speaking about myself here” (Foucault 1989a, 350).

So whilst it is true that Foucault engaged with, and responded to, the dominant traditions, it is fair to say that he saw himself as fundamentally orientated by his commitment to the tradition of French history and philosophy of science taught by Canguilhem (and through him, the philosopher of natural sciences Gaston Bachelard). This connection can be seen in his Introduction to Canguilhem’s ‘The Normal and the Pathological’, in which Foucault distinguishes this tradition from the dominant stream of French philosophy, in terms of how Husserl’s phenomenology was taken up in France.

\(^{24}\) I follow Gutting 1989, 103 n.11 in substituting “experience” for Sheridan’s “experiment” in translating the French “expérience”
Foucault asks “why … following its own logic, [the philosophy and history of science] turned out to be so profoundly tied to the present?” His response is that it “…avails itself of one of the themes which was introduced surreptitiously into late 18th century philosophy: for the first time rational thought was put in question not only as to its nature, its foundation, its powers and its rights, but also as to its history and its geography; as to its immediate past and its present reality” (Foucault in Canguilhem 1991, 9).

In other words, by viewing reason in terms of the contingent contexts which constitute it, French history and philosophy of science enabled an external standpoint from which to critique reason. It is this tradition that Foucault sees as taking up the Kantian project of the critique of reason not in terms of its nature, function or limits, but in a register that engages with the present by revealing its contingency. Foucault carried on this critique in the form of his critical histories of the human sciences which bore consequences for the contemporary world. Foucault sees this tradition of history and philosophy of science, extending back to Comte and transmitted through Bachelard and Canguilhem, as “serving to support the philosophical question of the Enlightenment” (Foucault in Canguilhem 1991, xi). This enables Foucault to place himself in historical relation to the Enlightenment. This is a theme I will turn to in Chapter 4 in my discussion of Foucault’s relationship to Kant, and his conceptions of critique and modernity. For now it is important to note that whilst Foucault engaged with the dominant philosophical fashions of his time, it was the philosophy and history of science that oriented his project.

However, this does not help us grasp the major commitments of Foucault’s work as a whole. The difficulty is that Foucault was constantly adjusting his theoretical positions, at times pronouncing bold provocations, at other times, revising or retracting previous positions. His ongoing retrospective interpretation of his work gave the sense of an experimenter, wanting to test ideas, to put them into circulation before committing further to them. In an interview in 1977, Foucault announced “What else was I talking about, in *Madness and Civilisation* or *The Birth of the Clinic*, but power? Yet I’m perfectly aware that I scarcely even used the word and never had such a field of analysis at my disposal” (PK, 115). Five years later he was saying, “it is not power, but the subject, which is the general theme of my research” (Rabinow & Dreyfus 1983, 209).

This was typical of the ongoing revisions of Foucault’s self-understanding and work. In some cases, Foucault deliberately bracketed questions for methodological reasons. In other cases, he seems to later recognise what may have been implicit earlier. However, the boldness of Foucault’s revisions and self-reinterpretations suggests a desire move on, to look back at his previous selves from new perspectives. In an interview in 1978 he said, “I’m an experimenter in the sense that I write in order to change myself and in order not to think the same thing as before” (EW3, 240). Foucault’s
reinterpretations of his work may offer clues, but it is ultimately for others to judge whether to or not to take him at his word. And clearly the conventional periodisation – archaeology, genealogy and ethics - gives rise to questions of consistency and coherence.

A number of interpretive schemas offer an overall sense and direction of Foucault’s project. I argue that none of these schemas is entirely satisfactory, since they neglect what Foucault himself thought about fitting into any general interpretive schema. In *The History of Sexuality* v 2, he sees philosophical activity as “the critical work that thought brings to bear upon itself … the endeavour to know to what extent it might be possible to think differently, instead of legitimating what is already known” (HS2, 9). Seen in these terms, Foucault’s philosophical project cannot easily be judged against fixed theoretical foundations. By squeezing his *oeuvre* into a general interpretive schema, we risk distorting the specificity of each of his works. Since his work does not aspire to complete systematicity, we must not place too much weight on harmonising his work within an overall schema by attempting to draw out a single set of philosophical methods and commitments that seem partially hidden. The danger is that such schemas typically interpret work in terms of the problems or questions of established disciplines. As Gutting argues, it is precisely established disciplines that are the problem for Foucault. To present Foucault’s work as within established paradigms is to deny, in advance, the thrust of his entire project to rethink the most fundamental presuppositions (Gutting 1994, 4).

Does this mean that Foucault was not a philosopher, but a clever savant who used philosophical language without serious commitment? Was Foucault a historian, rather than a philosopher? Gutting responds to Han’s claim that Foucault’s archaeologies slip into unintended transcendentalism by arguing that “The Birth of the Clinic and *The Order of Things* are both primarily works of history, not philosophy in the traditional sense”(Gutting, 2003). Whilst Gutting doesn’t tell us exactly what the “traditional sense” of philosophy is, it is certainly true that Foucault distances himself from certain understandings of philosophy. In particular, he distinguishes himself from philosophy understood as a socially disengaged, theoretical and universalist enterprise of seeking ultimate truths. Foucault did however, see himself as part of a stream of philosophical thought that has become more or less marginalised during the modern period. As I will discuss in Chapter 4, he becomes clearer about his place within this stream as his work progresses, eventually coming to identify philosophy as a way of life in which one is transformed, a form of *askesis* or spiritual exercise.

At the very least, Foucault’s project can legitimately be seen as an ongoing exploration which continues to challenge who we are today so as to change us and what we think. The central focus of

---

25 For example, Rabinow and Dreyfus (1983), Han (2002), Gutting (1989)
26 cf AK,19.
this ongoing experimentation in thought is the constitution of subjects, through the three axes of knowledge, power and subjectivity. In a 1983 interview, he tells us: “Three domains of genealogy are possible. First, an historical ontology of ourselves in relation to truth through which we constituted ourselves as subjects of knowledge; second, an historical ontology of ourselves in relation to a field of power through which we constitute ourselves as subjects acting on others; third, an historical ontology in relation to ethics through which we constitute ourselves as moral agents” (Rabinow & Dreyfus 1983, 237).

If we see these three domains mapping onto the three periods of Foucault’s work, we can say that these periods do not consist of a series of corrections on the path to a single approach, but that Foucault adopts whatever methods are required by a specific subject matter. Archaeology produces histories of thought which abstract from the individual subject in order to get at the linguistic structures which both constrain and make possible the fields in which individual subjects operate. Genealogy traces the origin of practices and institutions from their contingent beginnings, to display the hand of power. Ethics charts the possibility of resistance, whereby subjects are seen to be not totally determined by knowledge and power. The three methods are perspectives, each with their insights and blind spots.

It would be too tidy to claim that the works of archaeology, genealogy and ethics correspond precisely to the axes of truth, power and subjectivity. But it is fair to say that a central focus of Foucault’s work is the constitution of subjects in history, and although each of the axes – truth, power and subjectivity – is inextricably entangled with the others, each period of Foucault’s work thematises one of these axes in addressing the constitution of subjects. If we recall that science is the most authoritative form of knowledge in the modern era, we can see how science, power and subjectivity form a tangle that Foucault attempts to untangle in his genealogies. I am not saying that this is the only way to approach Foucault. But we can fruitfully see him firstly examining knowledge, then its link to power, then recognising the importance of self-constitution. With this approach all periods of Foucault’s work can be brought to bear productively on the topic of this work, science as a social institution.

3 HABERMAS’ LIFE, MOTIVATIONS AND COMMITMENTS

Habermas lived through the tumultuous years of war and defeat, the allied occupation and war crimes trials, the drafting of a new constitution and Germany’s social, economic and industrial reconstruction - all the time against the backdrop of the cold-war and the occasional threat of nuclear annihilation. As a child he had a series of operations on his cleft palate which “sharpened awareness of the deep dependence of one person on others” leading him to “those approaches that emphasise the
intersubjective constitution of the mind” (BNR, 14). At school his distorted pronunciation produced a strong interest in the conditions under which linguistic communication was successful, and the genesis of moral and social norms governing communal life (BNR, 16). As a teenager he had been drafted into the Hitler Youth Movement. After the war, 16 years old, he was appalled to not only to learn of Nazi atrocities, but to witness his elders making excuses for the perpetrators (Habermas 1983, 57). The young Habermas considered Heidegger his major influence, but in 1953 was shocked to find Heidegger’s republished 1935 lecture Introduction to Metaphysics which retained the original reference to “the inner truth and greatness of National Socialism” without any revision or commentary (BNR, 18). Responding with a newspaper article criticising Heidegger, Habermas took up the role that he has consistently maintained as a public intellectual engaged in political and social questions (Habermas, 1953).

This experience also motivated a search for conceptual resources beyond German thought, particularly within the Anglo-American pragmatist and democratic traditions. In 1956, Habermas became Adorno’s research assistant at the Institute for Social Research at Frankfurt. Here he encountered the Frankfurt School’s integration of social sciences with philosophy in the form of critical theory. In this period, Habermas started writing his Habilitationsschaft on the public sphere.27 Criticised by Horkheimer who saw him as too radical, he left the School in 1959 to seek supervision from Wolfgang Adendorf at Marburg and after gaining his Habilitation in 1961, took up the position of Professor of Philosophy in Heidelberg. In 1964 he returned to the University of Frankfurt to take up Horkheimer’s chair as the Professor of Philosophy and Sociology. From 1971 Habermas was a director of the Max Plank Institute at Starnberg, returning to Frankfurt in 1983 to teach philosophy. He retired in 1994 but taught in the US for several more years. He still maintains a prominent role as a public intellectual on matters to do with German foreign policy during the Bush years and more recently the politics of the EU crisis.

The most important intellectual influence on Habermas is the broadly conceived Marxist tradition that grew out of German idealism and inspired Horkheimer’s notion of critical theory. Much of the development of Western Marxism can be seen as a shift from “scientific” Marxism that concentrated on the productive base as the engine of change, towards a focus on the more superstructural elements of culture as autonomous processes. This “revisionism” involved bringing other thinkers and traditions to bear whilst still conserving a normative orientation. In this spirit, Habermas’ first major work, The Structural Transformation of the Public Sphere (hereafter STPS) represents a further development of, or arguably a break from, the earlier conception of critical theory. Like earlier critical theory, it provides an analysis - drawing together insights from history, sociology, literature,

27 Published in 1962 as The Structural Transformation of the Public Sphere.
psychology, economics and philosophy - in order to diagnose a situation and articulate emancipatory possibilities. More distant from the war and taking on additional critical resources, Habermas is better placed to develop critical theory as a response to the challenges of late capitalism.

STPS foregrounds the abiding commitments of Habermas’ career by charting the emergence, in the late 17th and 18th century, of a public sphere of discourse. By “public sphere”, Habermas means that within which “something approaching a public opinion can be formed” (cited in Edgar 2005, 31). This would include the exchanges of opinions that occur in face-to-face communication in public spaces such as lecture halls, salons, coffee houses and clubs, or through journals and newspapers, all of which enabled the formation of more rational opinions. The public sphere was both an institution and a politically effective idea, driven by the need for an arena in which matters of common concern could be rationally discussed and public policy critically assessed in terms of a concept of public interest. Although actually beset by contradictions, it nonetheless implies counterfactual norms which underpin democratic discourse. Habermas articulates this repressed normative core: Firstly, a type of discourse was conducted that “disregarded status altogether.” The authority of the better argument could assert itself against any form of status - rank, birth, prestige of public office or economic power. This idea, although seldom be actualised, is powerfully consequential as an ideal against which actual practice could be judged. Secondly, “the problematisation of areas that until then had not been questioned was possible.” Church and state authorities, previously holding a monopoly in interpretation in matters of public concern, could now be challenged. Thirdly, the public is seen as in principle inclusive. “However exclusive the public might be in any given instance, it could never close itself off entirely and become consolidated as a clique” (STPS, 37). These norms, applied to informal collective discourses, grounded claims of legitimate authority based solely on the rational claim of the best argument. Despite tensions induced by ideological distortions, the growth of the public sphere throughout the eighteenth century was accompanied by a series of reforms. Public opinion forged through argumentation, could now challenge the authority of power traditionally vested in the state. Throughout 18th century Britain, there was an increasing appeal to “the sense of the people,” “the common voice” and the “public spirit”. Such appeals represent the first moves by which parliament became responsive to the people, whose collective opinion legitimised parliamentary law-making (STPS, 31). Habermas’ entire project can be seen to rely on the derivation of the counterfactual normativity of the public sphere.

However, the bourgeois public sphere began to transform, firstly because those excluded challenged their exclusion from this supposed universal openness and reciprocity. From the time of the French revolution, the masses established themselves as political agents involved in class struggle who, later in the 19th century, formed trade unions to demand political and social rights. At the same time, capitalism’s boom-bust cycle increasingly stripped any appearance of natural justice from the free-
market and, from the beginning of the 20th century, the small-scale entrepreneur was displaced by larger companies. Increasingly throughout the century, the concentration of capital in large oligopolies undermined the notion of perfect competition. The state responded by expanding its activity and regulating the economy, becoming a major provider of goods and services. With economic rationalisation and technological development, the state and society became interlocked and the public sphere lost its role of go-between mediating between public and private. New larger powers increasingly came to contest and control this public space. After the First World War, government regulation and power increased in order to regulate capitalism as it moved from laissez faire to a more monopolistic form. The public sphere was transformed from a sphere of independent entrepreneurs, attempting to articulate their own standpoints, to a domain vulnerable to the vicissitudes of monopoly capitalism, where bureaucracies comprising business, government and unions could advance their own agendas in the political domain. These transformations, culminating in the welfare state, undermined the initiative and autonomy of the bourgeoisie, who came to see themselves as subservient to economic and government support and regulation. The vitality of the public sphere was thus diminished as autonomy was narrowed to the exercise of private preferences in leisure activities and life-style choices. With the rise of formally organised political parties, the politician became a professional, drawing upon advertising and public relations to elicit acclamation. The bourgeois public sphere, originally a conduit between society and the state, had assumed that rational debate could be free of economic interests and that all members of a universal public sphere shared common interests. Now in advanced capitalism, rational debate based on this commonality is replaced by the haggling and bargaining of special interest groups.

In many respects Habermas’ argument runs parallel to Adorno’s and Horkheimer’s 1944 *Dialectic of Enlightenment*. Like their critique of Enlightenment transformed back into myth, the bourgeois public sphere is transformed into manipulated publicity in which states and corporations secure plebiscitary acclamation. The difference is that Habermas does not abandon the Enlightenment ideals, but insists that the values articulated through the norms of the bourgeois public sphere have an enduring and functioning significance, both regulative and critical. Despite losses in discursive rationality, the idealisations embodied by these norms are not flawed. The decline of their normative power is best explained by changes in modern capitalist societies. For Habermas, the bourgeois public sphere remains a crucial normative dimension to understand democratic discourse.

Habermas’s critical theory differs from his predecessors in other ways. Since Adorno’s and Horkheimer’s version sought to enable individuals to resist homogenisation by current social reality, freedom was characterised only negatively by the capacity to resist. Habermas however links freedom inextricably to democracy and the renewal of its institutions and thus aims to identify social and institutional conditions in democratic institutions which foster autonomy. He develops a “quasi-
transcendental” account to underpin the progressive norms of critical theory. He doesn’t simply adopt Marx’s assumption that each stage of society would somehow “resolve” the contradictions of the previous stage. Instead of a speculative philosophy of history, Habermas turns to empirical sciences to reconstruct the irreversible and progressive stages of childhood learning which he extends as a developmental sequence explaining the stages of social evolution. On this account, social progress occurs in several interactive dimensions. The rationalisation of purposive-rational action (to do with growing mastery over nature) must be distinguished from rationalisation at communicative interaction (which enables political emancipation by removing restrictions to self-reflection and communication). This approach to critical theory provides standards by which current society can be judged, but cannot predict how a future society will necessarily develop. If the goals of society reside in free communication, we must reconstruct this in terms of social evolution in order to establish very abstract norms that transcend their embodiment in any factual society.

In Habermas earliest work, we see the commitment to democracy reflected in his articulation of the norms of the bourgeois public sphere and his interest in domination-free communication that would later form the core standards for his moral-political theory. In STPS, he already has the two-level model of democratic deliberation that he elaborates thirty years later in Between Facts and Norms (STPS, 244-48). This commitment to democracy and undistorted communication forms the background to Habermas’ various responses to science. After STPS, Habermas took critical theory in a different direction, still aimed at understanding the basis for democratic discourse, but now directed against the then dominant positivism, which he saw as a truncated form of reason distorting communication. This move must be seen in the context of the generally uncritical attitude towards scientific and technological progress prevalent in the early 1960’s, a time when the state played a major role in the economy, to the extent that “technocracy”, the idea that political decisions were merely technical and could be made by scientific experts, seemed a real possibility. Habermas responded by contributing to “the positivist dispute”, an exchange initiated by Adorno and Popper in which Habermas and Adorno defended a dialectical social science against the reductive tendencies of positivism (Habermas et al 1976, TP).

Habermas’ critique of positivism found its definitive formulation in Knowledge and Human Interests (hereafter KHI), which put forward a Kantian epistemology, whereby three deep-seated anthropological interests serve as quasi-transcendental conditions of three distinct domains of knowledge. Habermas surveys the “abandoned stages of reflection” to reconstruct how Kant’s

---

28 By “positivism” I refer to the philosophical attitude which restricts what can count as “knowledge” to what is manifest in sense experience, thus excluding value judgements and normative statements. See Chapter 5 for further discussion. Also see Kolakowski (1972)

29 Outhwaite cites Helmut Schelsky and Hans Freyer as associated with the technocracy thesis (Outhwaite 2009, 20).
differentiated notion of reason had given way to positivism’s view of the natural sciences as being the paradigm of knowledge *per se*. Drawing on Kant, Hegel, Marx, Peirce, Dilthey and Freud, this critique reflects on both the function of knowledge in the reproduction of social life and the historical conditions in which the subject of knowledge is shaped. It represents Habermas’ first attempt to provide a systematic framework for a critical social theory which he wanted to establish as a distinct form of knowledge.

As positivism lost its dominance, Habermas began developing a more comprehensive social theory drawing on various developmental accounts of morality and communication from the social sciences. The basis of his early concern with positivism – that it suppressed reflection and free communication - is sustained throughout his entire career, and is seen in his later engagement with Luhmann’s systems theory and more recent discussions of naturalism (LC, 130-42; BNR, 151-181). After KHI, his Kantian critical theory based on cognitive interests gave way to a theory of communicative action. This is a move against the “philosophy of consciousness” which relies on the conceptual dualism between subject and object. Like Heidegger, Habermas regards this as an abstraction from a more fundamental holistic, pre-reflective experience. Habermas had long insisted that reason resides at the very heart of human communication, which is the engine for human emancipation (KHI 314). By focussing on communicative structures, Habermas “post-metaphysical” theory of rationality and social integration, based on reconstructions of the competencies and normative presuppositions underlying communication, avoids any claim to ultimate foundations.

Habermas’ 1981 *magnum opus*, *The Theory of Communicative Action* (hereafter TCA) presents an alternative account of rationality by critically examining theories of rationality that informed Weber, Durkheim, Parsons and earlier critical theory. Rather than reason as the substantive concept of traditional philosophy, Habermas looks to a notion of formal or procedural reason. He moves critical theory away from the strong transcendental framework of KHI, towards a more modest, fallibilist, empirical account of the philosophical claim to universality and rationality. He will characterise this naturalistic approach as a form of “soft naturalism” whereby “weakly transcendental” conditions are directed to invariant structures and conditions and raise universal, but defeasible claims. Drawing on the speech act theory of Austin and Searle he shows how, in everyday life, speakers make claims that transcend specific conversational contexts. The expectations raised in such speech acts imply a set of rules that circumscribe the meaning of a domination-free dialogue undistorted by power. This provides Habermas with a rational standard against which he can discern various pathologies of modernity.

The 1980’s also marked a period in which Habermas engaged in debates about contemporary issues such as how Germany’s Nazi past should be remembered, immigration and neo-Nazism. His
commitment to a universal humanism led to an interest in cosmopolitanism and multiculturalism which played out in questions about the relationship between constitutional law and democracy, the role of the nation state and the “clash of civilisations” thesis. From his earliest work, Habermas was committed to the reconstruction of his own society as a liberal democracy capable of resisting totalitarianism. This continues with his later works on communicative action, legal philosophy, international law and post-secular society. In a 2005 interview he said “democracy is like a red thread through my work from the beginning until what I am presently interested in” (Habermas 2007a). Criticising all forms of fundamentalism, Habermas reconsiders the relationship between reason and religion which, in TCA, he had treated from a sociological perspective, as an archaic mode of integration. Now he will explore the relations between religious and philosophical modes of discourse and the role of religion in politics. He grants that religious discourse still harbours potentials of meaning from which philosophy can learn (RR 77, 162). And with the positivism of the 1960’s long discredited, Habermas now defends the “incomplete project of modernity” and its rational foundations against anti-modernists, traditionalists, conservatives and “postmodernists” (with whom he includes Foucault).

Whilst conscious of philosophy’s ultimately broader purview, Habermas recognises the emancipatory potential of science and incorporates it into his theorising to provide perspectives not available to philosophy. He is however concerned that science increasingly links itself to systemic imperatives such as the economy, overstepping its proper domain, to short-circuit democratic deliberations. For example he argues that technologies such as human cloning are advancing so rapidly and are so tied to economic imperatives that they bypass rational scrutiny and informed public discussion (FHN, 18).

In his theory of deliberative democracy, set out in BFN (1992), we see Habermas working out the issues first raised in STPS, bringing his highly complex and idealised discourse theory to bear on the political-legal institutions of complex modern societies. In one sense Habermas is staunchly Hegelian, locating his theory’s normative dimension in existing institutions. Unlike Foucault, who sees Kant as the founder of modernity, Habermas thinks it was Hegel who first saw modernity as a philosophical problem. Whilst distancing himself from the excesses of Hegelian metaphysics, Habermas seeks to recover the hidden traces of immanent reason so as to provide a better grounding for a critical social theory of modernity. Philosophy has come to recognise that consciousness has no unmediated access to being and the subject is conditioned by a multiplicity of historical, economic, social, psychological and cultural factors. It must therefore direct its attention to formal conditions of rationality, such as those found in language.

Habermas critical philosophical theory is not distinctive because it endorses a particular scientific theory or method. In fact, he uses social sciences in methodologically and theoretically pluralistic
ways. What is distinctive about critical philosophy is that it unites normative and empirical inquiry. The motivating commitment that lies behind all Habermas work is the idea of human emancipation in terms of the establishment and maintenance of democratic institutions which require interactive open-ended discourse free as far as possible from all constraints, including the internal constraints that Kant referred to as “immaturity” (PT, 29). Despite the many directions that Habermas’ practical engagements and sources of inspiration have taken him, he has never strayed from this commitment.

Yet within in this vast body of work, science is always crucial— in two respects. Firstly we see in his criticisms of positivism and reductive naturalism that what he criticises is not the content of science, its truth or falsity, but a philosophical misunderstanding of science that awards it the status of the most authoritative form of knowledge, which can be unproblematically utilised for practical ends. This is the basis of his critiques not only of positivism, but also systems theory, technocracy, genetic engineering, neuroscience, economic rationalism, and many more areas. Secondly, in these very critiques, Habermas harnesses sciences such as linguistics, anthropology, sociology, psychology as resources that enable an “external” perspective, capable of being harnessed as a normative standpoint, to cut across conditioned philosophical reason.

4 PHILOSOPHY, SCIENCE AND MODERNITY

Both Foucault and Habermas thought that philosophy’s relation to science was vitally important. They shared a perspective which sees philosophy as providing a more comprehensive view by virtue of its self-reflexivity. Consistent with the Critical Theory tradition, both see philosophy as accepting, addressing and harnessing the insights of the sciences whilst still offering a critical perspective by drawing on a broader context oriented to emancipation. This common outlook contrasts to certain contemporary attitudes, paradigmatically Quine’s, which tends to view philosophy as simply the more speculative end of science. Like Quine, both Foucault and Habermas see philosophy as needing to give up all pretensions of playing a foundational role. But unlike Quine, both recognise that it offers a broader purview. Foucault would agree with Habermas’s account that philosophy crosses different value spheres and domains of knowledge to bring highly specialised scientific knowledges back into the shared lifeworld where it interprets their significance and shows their limitations. Fundamental questions about the human condition cannot be answered by purely social or natural scientific approaches. Habermas emphasises that philosophy must develop a co-operative relationship with social sciences and empirical disciplines in general (TJ 286). It can’t think of itself separately from science, since it is the inter-relation between science and the lifeworld that provides it philosophy’s problems.
Foucault employs his genealogical critique more radically, to foreground science as a contingent human activity embodied in multiply-conditioned individuals within historical societies, so revealing the limitations of the view of scientific reason as reason per se. For Foucault, philosophy is a “spiritual” activity that transforms the subject and hence discloses the world differently. \(^{30}\) He traces the history of a subterranean tradition within philosophy that doesn’t seek absolute foundations, and doesn’t issue in universal propositions or scientific theories, but starts in the middle of things as a historically situated enquiry directed at the present situation. Like Habermas, Foucault’s perspective, linked to the insights of Critical Theory, sees itself as engaging with the sciences not theoretically, but in a way which orients everyday life.

But there are also differences. Habermas seeks “the reconciliation of a modernity that has fallen apart, the idea that without surrendering the differentiations which have made modernity possible … one can walk tall in a collectivity that does not have the dubious quality of backward-looking substantial forms of community” (Habermas 1986, 125). Modernity is the “incomplete project” which must not be abandoned (Habermas 1996). He wants to endorse cultural modernity by taking its knowledge and interpreting it for contemporary society, so showing not only the limitations of science but, something Foucault barely canvasses, its emancipatory potential. He offers a developmental account of progress, whereby the sciences are linked to cognitive advances of distanciation and differentiation. But it is a progress that is by no means guaranteed or universal and comes at considerable cost and risk. Foucault is more wary in his assessment of modernity, especially the human sciences and their ubiquitous power effects. Yet both thinkers are oriented to the present situation or crisis. Both draw on broad historical contextualisations to open up a critical perspective enabling us to see our categories of thought as not given by nature, but as historically conditioned and contingent, and therefore subject to change.

5 AIM

Some of the specific questions that Foucault’s work raises include: Are there broader formations of knowledge that cross boundaries between science and non-science which can be thought of in terms of structures enabling and constraining what can be said and thought in a particular era? How does power actually operate and how is it entwined with scientific knowledge? In what sense can man be an object of science? Is there an essential human nature to be known by science which, at the same time, is known by virtue of that nature? How is power resisted and how does philosophy relate to political action. Habermas raises an overlapping set of questions. But there are important

\(^{30}\) Foucault uses, “spirituality” free from any religious connotations (HS, 15).
methodological differences. For example, Habermas locates his analysis in a quasi-transcendental framework, supported by counterfactual idealisations such as “the ideal speech situation”\(^{31}\). Foucault, insisting on concrete practices and specific events, refuses such idealisations. Yet both philosophers, critical of each other, recognised the limitations of science and challenged its status as the dominant paradigm of knowledge in modernity. For both thinkers this is a political project in which they seek to reveal consciousness as conditioned by a vast range of historical and social factors, in order to change consciousness. Both recognised the need for a more reflexive perspective by which philosophy could articulate social problems not visible from the specialised perspectives of science.

My task is to tease out points of difference between Foucault and Habermas and, without smoothing away those differences, draw them into dialogue, to reveal the relations between science and society in their complexity, with all the attendant risks, dangers and opportunities. By this engagement, I will gradually tease out the conjecture of a broader framework, which reveals the positions of Foucault and Habermas as two distinct tendencies within the structure of modern thought. These positions are revealed not so much by what Foucault and Habermas say directly about science, but how they use it, the role they grant science in their own critique of reason. My approach is informed by the hermeneutic circle. The back and forth movement between parts and wholes is extended beyond statements, texts and œuvres to the biographies and contemporary and historical engagements I have sketched. Whilst the hermeneutic circle informs my approach, it doesn’t prescribe a form to this work, which for clarity of presentation adopts a chronological account of the development of firstly, Foucault’s and then Habermas’ thought before bringing both together in a concluding chapter.

A major animus of my work is to support the general thesis that knowledge, truth and rationality must be granted an authority that extends beyond the forms they take within the sciences. This theme, opposed to a long thread of reductive thought that runs from positivism to contemporary forms of reductive naturalism, is found throughout Foucault’s and Habermas’ work. My aim is to show how this broader standpoint is able to reveal not only the contingent constitution of science but its social effects, in terms such as power, lifeworld colonisation and cultural impoverishment. By contextualising science as an object within a broader setting, we can see how it is put together in particular social, political and historical contexts, and ask what role it plays within those contexts and how it could be otherwise.

The following three chapters will deal with Foucault’s work in terms of the three conventional periodisations I have discussed. Chapter 2 will consider the archaeological method, as a form of historical analysis of scientific knowledge. I will discuss The History of Madness which problematises

\(^{31}\) See OPC
the concept of madness in a way that will be picked up as a critique of the psychiatric and psychological sciences in Foucault’s later lectures. I will consider *The Birth of the Clinic* by highlighting the importance of medicine in setting up normative frameworks for the social sciences and foreshadowing the later development of Foucault’s concept of “biopower”. I will then discuss *The Order of Things* as Foucault’s major engagement with Kant’s critical philosophy during this period and what Foucault considers to be its problematic legacy in the social sciences. I will finally consider Foucault’s analyses of the empirical and human sciences, and his three structuralist “counter-sciences”. Chapter 3 will chart Foucault’s shift to genealogy by his engagement with Nietzsche, and the subsequent thematisation of power in *Discipline and Punish* and *The History of Sexuality* vol 1 and the *College de France* lectures. I will discuss the constitution of subjects, objects and scientific concepts, and the emergence of new population-based sciences, for which Foucault develops the concepts of biopower and governmentality. Chapter 4 will chart Foucault’s shift of focus from power to the subject through his reading of ancient philosophy, and his adoption of the concepts of care of self and spirituality. I will discuss how Foucault’s history of philosophy enables him to consolidate his work by drawing on his earlier phases to place himself within a tradition of philosophy. Drawing on his *History of Sexuality* vols 2 & 3, interviews and the last four years of lectures at the *College de France*, I will trace Foucault’s genealogy of the subject from antiquity, its constitution by pre-scientific knowledge and its reconstitution as it emerges into modernity and modern science.

Chapters 5, 6 and 7 will more or less chronologically chart the development of Habermas’ theoretical framework and engagement with a range of interlocutors and issues in relation to science. Chapter 5 will offer an exposition of Habermas fundamental attitude to science in the positivist dispute and then discuss the early formulations of Habermas’ theoretical framework in *Knowledge and Human Interests*. Chapter 6 will consider Habermas’ linguistic turn and his developmental theory of rationality within the system-lifeworld paradigm of his social theory set out in *The Theory of Communicative Action*. I will address the richer articulation of the social role of science and the pathologies of modernity that this theory enables. Chapter 7 will consider a number of more applied aspects of Habermas’ understanding of science and society, including the notions of freedom and determinism, deliberative democracy, genetic intervention and the place of religious discourse. I will conclude by considering how Habermas sees the relation of philosophy to science. Chapter 8 will bring Habermas and Foucault together to tease out the conjecture of a broader framework, which reveals the positions of Foucault and Habermas as two tendencies within the structure of modern thought.
Chapter 2
Archaeology

By 1960, De Gaulle had been in power two years and France was modernising, exploding its first atomic bomb in February. However, modernism did not extend to all aspects of French culture. In July, homosexual acts were outlawed by a statute that remained in force until 1981 (Macey, 91). Foucault met Daniel Defert, a committed campaigner against the Algerian War, in September. He remained Foucault’s partner. Whilst Defert gave Foucault a taste of political militancy, Foucault continued to concentrate on his literary and intellectual pursuits (Macey, 93). At this time, he was primarily concerned with publishing his doctoral thesis which would give him a place in the French university system. The thesis was published in 1961 as The History of Madness (hereafter HM). But Foucault was disappointed by its critical reception (Szakoczai, 206). His exploration of relationships between power and knowledge were greeted with silence from the intellectual left. With a strong sense of being misunderstood, Foucault had only vague notions about the work he intended to pursue.

By the late 1960’s the situation had changed completely. Foucault now had several books behind him. In 1966, The Order of Things (hereafter OT) immediately became a best seller. He relished the success. But soon he became disillusioned with media excesses surrounding star status. In 1969, The Archaeology of Knowledge (hereafter AK), was written as retrospective overview to order and clarify his archaeological method. During the 1960’s, we can see Foucault gaining clarity about his project and increasingly distinguishing himself from the intellectual influences I discussed in Chapter 1.

In the years before his death in 1984, Foucault frequently proclaimed a fundamental unity to his intellectual project. He presented systematic classifications of his works, which he saw as revealing the contingent historical construction of the subject along three axes – knowledge, power and ethics (EW1, 262-3, 318; EW3, 326-7; HS2, 6; FR, 336-8). This chapter examines archaeology, or the knowledge axis, specifically the human sciences. It describes a fundamental level concerned with the formation and transformation of objects, concepts, forms of cognitive authority and social function, which forms an unconscious and anonymous structure of constraints and possibilities underlying what can count as scientific knowledge.

I will firstly compare archaeology with more orthodox approaches to the history of science. Whilst Foucault’s work varied from his attempt at clarification in AK, this book established a framework to consider archaeology in relation to conventional histories of science. I will then build on this framework, starting with Foucault’s first book HM, an ambitious, vast, rambling work that anticipates
all his later themes and preoccupations. Discourse, power, practices, institutions, genealogy and ethics are all found in HM. I will then turn to *The Birth of the Clinic* (hereafter BC) published in 1963. The broader significance of BC is that it demonstrates the seminal role medicine played in the formation of the human sciences. Finally I will examine OT to consider the place of the sciences, within the modern *episteme*. 32 Here Foucault argues that we are entrapped in an “anthropological sleep” by our uncritical acceptance of a concept of Man, initiated by Kant’s transcendental philosophy. I will conclude the chapter by examining the Death of Man thesis as a critique of human sciences as well as addressing some of the concerns that Foucault’s archaeology has raised.

1 SCIENCE, HISTORY AND ARCHAEOLOGY

Foucault’s archaeology extends the French tradition of history and philosophy of science to specifically deal with the complexities of the human sciences. Like Bachelard, Foucault’s histories of science comprises various threads within different regions of scientific work which couldn’t be seen as a unified development of rationality (AK, 4). But like Canguilhem, Foucault softens Bachelard’s epistemological ruptures, allowing a degree of continuity whereby continuous concepts are displaced and transformed within discontinuous theories (AK, 5). 33 Scientific concepts are not simply parts of theories used to interpret phenomena, but provide the cognitive and perceptual frameworks for theoretically grasping phenomena. Foucault’s archaeology extends Canguilhem’s “history of the concept”, from particular disciplines to the analyses of *epistemes* that ground the possibility of several disciplines. Archealogy excavates basic organising concepts, cutting across apparently discrete disciplines to determine the nature and extent of possible knowledge within the *episteme*.

Bachelard and Canguilhem were dealing with established scientific disciplines, and hence could adopt the normative standpoint of current science. However, since Foucault’s archaeology examines discourses which have not reached, and may never reach, the threshold of the sciences, his archaeology can’t assume the normative stance of current science. Within human sciences such as psychiatry or criminology, with objects such as madness and criminality, the norms of the best current science cannot be used to discern the scientific from the non-scientific. Archaeology uncovers discursive practices that underlie a corpus of knowledge that aspires to the status and role of science

---

32 Foucault employs the term *episteme* to refer to the unconscious structures which underlie the production of scientific knowledge in particular societies. (OT p.xxiii) (AK 211)

33 Bachelard characterised epistemological ruptures as decisive counter-intuitive breaks of scientific knowledge from common sense and previous scientific theories requiring new concepts, for example the development of quantum theory (Gutting 1989, 14-5).
(AK, 210). It reveals an underside of reason, obscured by orthodox historical accounts of the history of science which ideologically presents the inevitable growth of reason towards the present.

Foucault’s normative agnosticism and emphasis on discontinuity would be familiar to readers of Thomas Kuhn (Kuhn, 1962). There are however, instructive differences. Kuhn’s discontinuity is embodied in his notion of scientific “revolutions” which start with crises. After reaching a climax, the crisis subsides and the scientific community re-establishes consensus which ushers in “normal” science to solve problems within the bounds of the new paradigm. In contrast to Kuhn’s emphasis on consensus, Foucault looks for unarticulated pre-conceptual structures that regulate normal science. Rather than socially-achieved paradigm shifts explaining the continuity or discontinuity of theories, Foucault remains at the level of underlying structures of which we remain unaware.

Whilst Kuhn draws on his model to explain why science is both progressive, yet not as rational as we are tempted to think, Foucault employs his notion of discontinuity strategically to undermine the privileged role of the human subject. This enables him to get at the structures which in some sense determine our thinking and speech, yet are beyond the grasp of thinking and speech. He wants to bring to light how the science which arises from these structures comes to dominate and shape us. For Foucault, it is the transformation of concepts not subject to conscious awareness and control that makes possible the development of individual scientific disciplines. Since this history of transformations is not visible from individual disciplines, the archaeologist must go to a level of conceptual history more fundamental than first-order concepts of individual scientific disciplines to reveal the concepts which define the possibility of these first-order concepts. Foucault can’t simply align divisions within his analyses with the intentional products of human subjects - essays, books, or the oeuvres of particular authors. Nor can he assume that works of different authors are related by means such as transmission, influence, tradition or the “spirit of the age”. Foucault’s groupings reflect the deeper movement of knowledge within “discursive formations”, resulting in blurring of boundaries between, and new divisions within, conventional groupings (AK, 32). Discursive formations do not describe particular disciplines, scientific or otherwise. The discursive formation from which a science emerges, extends beyond science which remains connected to other knowledges within the discursive formation (AK, 203). Thus Foucault, unlike phenomenologists, doesn’t think there is a clear separation between the second-order idealisations of scientific discourses and what

---

34 Hacking provides a good, though preliminary, comparison between the uses of the concept of discontinuity found in Kuhn and Foucault. My comments are to some extent an elaboration of Hacking’s thoughts (Hacking 2002, 87-9).

35 Archaeology operates at the level of “statements” - functions with no reality unless inserted in a rule-governed system, in which they perform this function. Statements are united into “discursive formations” by rules that govern their formation in relation to their elements, such as objects, cognitive status, concepts. (AK, 89-98)
belongs purely to the lifeworld. The boundary separating the two remains blurred. Scientific idealisations construct everyday experience by constituting subjects and objects, as part of the ordinary lifeworld. Rather than the lifeworld serving as the fundamental level underpinning particular scientific theories, Foucault posits an archaeological level which structures not only scientific discourses but also subjective experiences.

Since there are generally strong links between science and political, economic and religious ideologies within the discursive formation of a particular era, ideology is a more or less natural accompaniment to science, as one way in which science functions in relation to other knowledge (AK, 204). Ideology doesn’t necessarily exclude scientificity (AK, 205). Ideological bias and scientific objectivity are two entwined threads arising from the location of several discourses in a common discursive formation. Because the archaeological level of knowledge is not open to conscious scrutiny, but requires analysis, this entwinement cannot be untangled easily. Science serves as ideology “in so far as science, without being identified with knowledge, but without either effacing or excluding it, is localised in it, structures certain of its objects, systematises certain of its enunciations, formalises certain of its concepts and strategies” (AK, 204). Foucault is not suggesting that the mere content of science, bearing on a limited domain of reality, is ideological. Rather scientific discourses restructure other discourses within the discursive formation beyond science’s strictly limited domain. An example might be the slide from biological evolution to social theory. Here it would seem that the epistemic authority of science (which is only earned by being open to challenge) is subsumed into unchallengeable background knowledge (savoir) which it restructures.

As we will see in Chapter 4, Foucault links the French tradition of the philosophy and history of science to Enlightenment critique, through Kant, Hegel, Marx and the Frankfurt School. This tradition questions rational thought, not simply its nature, foundations or limits, but the contexts in which it has emerged and its role in the present (NP, 9). It is concerned not only with the internal questions of truth, but is a critique that brings the specificity of historical contexts in which that truth arises to bear on the contemporary world. Foucault’s archaeology is history in a sense that demands elaboration. Its basic premise is that during any given period, there is a shared unconscious framework which both constrains and enables thought. This framework - variously referred to the historical a priori, the episteme, the archive - structures observations, discussions and concepts by ordering and determining what can appear as objects of knowledge.36 By analysing these structures and their histories, we come to recognise the contingency of our own thinking and the possibility of thinking differently. Foucault envisages a history of things commonly thought to be outside history, such as reason or the subject.

36 Foucault introduces the term “historical a priori” to express the role of discursive formations as conditioning the thought that goes on within them, whilst at the same time being subject to historical shifts. (AK, 143-4)
By historicising these pre-supposed universal structures, thought to underlie history’s multiplicity and contingency, Foucault interrogates the present (BB, 3). By suspending commitment to the universals of the human sciences, he requires explanation of what is assumed.

Foucault’s accounts of change are layered and complex. By excising any suggestion of progress, they reveal the movement of history in its own right, uncluttered by teleology. He decentres the subject from its traditional explanatory role in history and seeks to grasp relations and regularities of which actors remain unaware (AK, 31-2). He eschews straight-forward causal accounts of how politics or economics determines consciousness. The archaeological method is structuralist in the sense that it simply describes the theoretical coherence of discourses among themselves in a given period (EW2, 285). It does not imply epistemic determinacy. The development of science is always unpredictable. In OT, Foucault entirely ignores non-discursive practices. He treats discourse as independent and autonomous to highlight the relationship between discourse and knowledge.

Archaeology brackets subjective intentions, truth-values, and even whether or not statements make sense outside their context (AK, 7; RD, 49). This is seen in the opening pages of BC, where Foucault recalls without comment the bizarre account of the 18th century doctor Pomme who “treated and cured” a hysteric by giving baths, ten to twelve hours a day, for ten months, and observed the internal organs expelling themselves (BC, ix). This account is one part of a before and after image that Foucault typically employs to highlight difference between either side of an epistemological rupture. Such differences are radically incommensurable. There is no common measure between Pomme’s account and our experience. Whole bodies of discourses are incommensurable because the systems of possibility underlying what counts as candidates for truth are completely at odds. These differences disrupt the smooth narrative of accumulated knowledge and reveal history, including our present, as open to alternative possibilities.

Archaeology doesn’t naively attempt to reconstruct the past “as it really was” but rather, consciously and self-reflectively assembles, regiments, orders and organises its traces. This doesn’t mean that “anything goes” or that history is merely a function of the archaeologist’s needs. Whilst archaeological accounts are directed by the archaeologist’s concerns, they are directed towards, and constrained by, a recalcitrant historical reality regardless of those concerns. This reality, though ordered by the activity of the archaeologist, is not reducible to a function of that ordering, but always exceeds it. Concepts of madness or criminality are not simply socially constructed, but have a real material basis in things such as physical bodies and behaviour (in the sense of bodily movement). Social objects emerge in history through the relations and interactions between their material bases, concepts, institutions and social practices.
Foucault wants to write “a history of the present” (DP, 31). He wants to explicitly and self-reflectively organise the past in a way which highlights the **historicity of the present**, the very contingency of the construction of what appears naturally given. Foucault’s “present” refers to practices that are current yet their constitution is linked to the past in ways we don’t realise. A “history of the present” lays bare that constitution and its consequences in a way that opens up freedom for change. Whilst conventional history explains away by smoothing over difference, Foucault dramatises it in order to throw the present into relief as a historically specific form of rationality.

**2 MADNESS**

I will now turn to Foucault’s first book, HM. If we recall Foucault’s biography – son of surgeon, student of Canguilhem, qualified in psychology, working in mental asylums, homosexual when homosexuality was considered a disease, suicidal thoughts, depression – it is not surprising that his first book problematised psychiatry and psychology. Nor is it surprising that Foucault remained engaged with the question of marginality, seeking to understand, and resist the pervasive power of science.

HM charts the history of the *experience* of the different ways madness has been socially constructed over three *epistemes*. In the Renaissance, madness remained in dialogue with, and was acknowledged by, reason. The pauper was related to the suffering Christ, the madman to the madness of the cross. In the middle of the 17th century, “a decisive event”, the Great Confinement, dramatically changed the experience of madness as its meanings were transformed, silencing the dialogue between reason and madness (HM, 77). By a restructuring of moral categories, poverty became an object of condemnation and idleness rebellion. Amongst those confined, the characteristics of what had been discrete groups became homogenised within the broad social category of Unreason (HM, 82). Here we see the *constitution* of a new experience of madness, by which “something inside man was placed outside of himself, and pushed over the edge of our horizon ... creat[ing] alienation” (HM, 80). This gesture was decisive in initiating the objectification of a part of man that had been an experience of inalienable interiority.

Mental illness became an object of scientific inquiry in the late 18th century, as confinement gave way to psychiatry and psychology. Whilst asylums granted some freedom, they imposed a crushing psychological confinement, by which bourgeois morality manipulated the mad to feel guilty. Again

---

37 Reflecting distrust in the categories of scientific psychology, Foucault doesn’t refer to “mental illness” as a fact that needs to be explained. He criticizes historians who suppose “an immutable continuity in madness.” (HM 79).
According to Foucault, the doctor’s entry into the asylum was not due to his scientific knowledge but his moral authority. He was seen as almost a magician, a worker of miracles. Foucault notes the paradox of seeing “medical practice enter the uncertain domain of the quasi-miraculous just as the science of mental illness was trying to assume a sense of positivity.” On one hand “madness is placed at a distance in an objective field where the threats of unreason disappear”. On the other hand, “the madman and the doctor begin to form a strange sort of couple, an undivided unity where complicity is forged along very ancient lines” (HM, 507). Foucault is referring to the easy slide into the uncritical idealisation of the psychiatrist and his knowledge, a slide pointing to epistemic stances that can develop towards authority of both science and superstition (HM, 509).

The resultant form of objectivity was a “magical reification” in which both doctor and patient were complicit, a myth of scientific objectivity that served to disguise moral domination in the name of bourgeois values. Doctors, refusing to recognize “the ancient powers that lent their status full strength” and having no other way to explain their power to heal, could only see themselves as searching for objective truths (HM, 509). Whilst 19th century psychoanalysis listened to the madman, this was not an engagement, but a monologue “that exhausted itself in the silence of others … the patient was trapped in a relation to the self that was of the order of guilt” (HM, 496-7). It was only with Freud that the doctor-patient relationship was itself highlighted as a scientific object. Yet whilst the magical effects of the relationship was granted its true importance, Freud’s explanation was still covered over by further myths of scientism, leaving a structure of moral judgement and coercion (HM, 510-11).

Gutting points to certain passages of HM, in which Foucault is drawn to the idea that “the voice of madness” itself can take us beyond the confines of rational categories. (Gutting 1989, 96-9). I think it more likely that Foucault was referring to the evocation of madness in the lyricism of artists and poets - Holderlin, Nerval, Nietzsche and Artaud - which might serve to reinvigorate the forgotten dialogue between madness and reason (HM, 518). As Rajchman points out, such a reading is consistent with a 1960’s tendency that saw revolution emerging from avant-garde writing. Foucault saw literature as a counter-discourse that transgresses limits, making them visible and contestable (Rajchman 1985, 111). Here we see the motivation that ran through Foucault’s entire project. Whilst archaeology marks the limits of what can be said, thought or experienced in an age, it also implicitly reveals what may
fall outside these limits. A domain of freedom opens up in literature which is capable of revealing different experiences based on perceptual and practical grids that differ from conventional science and rationality. This realm of freedom was articulated in Foucault’s later work as experimental self-transformation. These avant-garde writers disrupt the sharp conceptual distinctions on which rational thought relies, showing that “in man, the interior is also the exterior, that the extremity of subjectivity blended into the immediate fascination of the object, that any ending was the promise of an obstinate return” (OT, 518). On my reading, Foucault is gesturing towards the “doubles” that will become thematic in OT (which I will discuss shortly). The voices of madness evoked by these works reveal the tensions within the concept of man as both a subject and object of knowledge. In psychology, these same truths are revealed, but here reflective thought “protect[s] itself, affirming with growing insistence that the mad were nothing but objects, medical things” (HM, 519). Viewed through the lens of scientific objectivity, these truths split into irreconcilable antimonies, resulting in conflicting interpretations of madness within psychology (HM, 520-1).

Another interpretation sees HM as the account of exclusion, starting with the Great Confinement and leading to more subtle exclusions in the asylum. This is a picture of modernity in which power and rationality exclude, banish or subjugate freedom and madness. Koopman suggests that this reading assimilates Foucault to Weber’s theory of modernity as an age of relentless rationalisation, bureaucratisation and the categorical differentiation of rationalised value spheres (Koopman 2002). Whilst it is undeniable that HM is deeply ambiguous, I agree with Koopman that more fruitful readings are justified. Koopman thinks Foucault is more interested in negotiating the tensions between couples like power and freedom, rationality and madness. If modernity is based on a logic of exclusion the remedy would be liberation, whereby madness was liberated from reason, freedom from power. Yet as we will see with the “repressive hypothesis” in chapter 3, Foucault thought liberation was simply a enslavement to a new power. This is because madness and rationality, freedom and power presuppose each other. “Man and madman are bound by an impalpable connection of truth that is both reciprocal and incompatible” (HM, 529). Koopman sees purification as the logic of scientific modernity in which reason must preserve madness as its other. Purification requires that rationality and freedom are unmixed with madness and power. Purification is a process in which two concepts rely on each other such that the practices and objects they pick out must rigorously isolate themselves from one another. Purification of madness and reason amounts to the simultaneous production of both madness and reason in such a way that neither can admit admixture with the other.

Rather than liberating madness from its exclusion, Foucault wants to re-establish the dialogue between madness and non-madness. This involves transgression and experimentation. Foucault

38 For example, Gutting describes the Classical experience of madness in terms of “rigorous exclusion”. (Gutting 1989, 83)
laments that “compared to the incessant dialogue of reason and madness during the Renaissance, classical internment had been a silencing” (HM, 496). He wants to restore that interaction. He commends psychoanalysis for recognising the primacy of dialogue, but castigates it for turning it into a monologue. Foucault laments that Nietzsche can no longer be on the border of reason and unreason since that border is precisely what modern purification rejects (HM 537).

From the time the madman was taken into the asylum, he was alienated, becoming an object in the eyes of others and an object in his own eyes, a deep, inaccessible and problematic object. This account is not only about the madman, since “to recognise the mad was to recognise oneself, feel the same forces, hear the same voices and see the same strange lights rise up within.” The scientific gaze that objectified “could no longer see without seeing itself” (HM, 519). It is by another, who we recognise as a subject, but a subject totally objectified, that not just the madman, but man becomes an object to himself, fully open to scientific investigation, like any natural object (HM, 525). Such self-objectification requires a historically particular stance. This is perhaps what Foucault thinks madness can show us in the lyricism of poets in which opposites play. When we vacillate between seeing the madman as object and recognising him as subject, or as mad or non-mad, we see how one perspective relies on the other to constitute its objects.

3 THE BIRTH OF THE CLINIC

BC is Foucault’s most structuralist work. However, whilst Foucault was attracted to the anti-subjectivist stance structuralism offered, archaeology is distinguished by the historical nature of its structures. BC is an epistemological critique of medical knowledge attacking positivist understandings of scientific method as involving mere observation of what is immediately apparent. It reveals the interweaving of the perceptual and the discursive, in a series of adjustments accompanying the emergence of modern medicine. Its broader significance lies in showing the seminal role medicine played in the formation of the human sciences and anticipating the theme of biopower developed in Foucault’s later work.39

The epistemic shift to modern medicine must be seen against the background of Classical medicine which situated diseases within an ideal schema, defining their natures in terms of a priori essences, which it sought to discern by their manifestations within particular individual bodies. Rather than being regarded causally, diseases were seen to be related by qualitative resemblances, enabling tabular

39 Biopower refers to power directed towards the enhancement of life at the level of the population. (see Chapter 3)
formulations (BC, 10). This Classical picture came under challenge from medicine associated with epidemics which, by identifying causal processes by which diseases spread, provided a precedent for a new clinical medicine (BC, 23). In the late 18th century, a new mode of observation emerged which involved many observers pooling results of many reported phenomena. Probability enabled a new way of considering the range of observer perspectives on, and phenomenal variations within, a disease (BC, 102). Medical knowledge was constantly revised to establish causal links between diseases that had nothing in common according to the old classification (BC, 26). The clinic constituted a new interpretive grid by comparisons of phenomena with normal organic functioning, frequencies of occurrence etc (BC, 95).

BC examines those practices which enable human beings to treat themselves as objects in the purest sense. The “gaze” - a mode of perception that reveals hidden truths by examination - gained central importance as a mode of observation. Human beings were now able to place themselves at a distance, as objects of scientific investigation. Suddenly doctors were able to see and say things that previously were neither visible nor expressible. If knowledge is to be more than a report of a cluster of perceptual data, experience must be guided by the categorical framework of language, whilst language must be constrained by perceptual experience. Language guides what our visual attention is drawn to, just as it is expanded by what initially resists it. The clinic involved purging old theories by a perceptual stance which silenced them, overturning the old Classical codes which had determined what was seen (BC, 107).

These adjustments between perception and language are prompted by “the silent configuration in which language finds support” (BC, xi). This includes the pragmatic context in which language and perception are employed. It particularly refers to the archaeological level, in which “medical objects, perceptions and concepts” are reconstituted through a new experience into a new coherent, unitary model (BC 196, 51).40 The mutual adjustments between what can be seen and what can be said, against “the silent configuration in which language finds support”, reflect changes at the archaeological level that include epistemological, ontological and conceptual changes to the fundamental structures enabling and constraining discourse within a domain.

Further development of the concept of disease came about as the clinic was transformed by integrating anatomical dissection into its knowledge and experience (BC, 129). This change reflected fundamental and far-reaching shifts in epistemology. Rather than being concerned with only directly observable phenomena, temporal patterns and statistical frequencies, attention became focussed on the

40 These three elements – objects, perceptions and concepts – are elements of statements bound by rules which constitute discursive formations (AK, 34-43).
lesion, the precise anatomical site of a disease (BC, 126-36). Medical diagnoses still involved classification, but rather than a primitive fact, classification depended on the anatomically specific site of the lesion. The gaze was further transformed, “no longer merely read[ing] the visible” as reality, but also freeing its “implicit structures” to “discover its secrets” (BC, 120). A “fine sensibility”, developed from the doctor’s erudition, training and experience, is now employed. Rather than a gaze which “implies an open field, and … the successive order of reading”, Foucault refers to the glance which “strikes at one point, which is central or decisive [and] chooses that line that instantly distinguishes the essential; it therefore goes beyond what it sees. It is not misled by the immediate forms of the sensible …” (BC, 121).

With the establishment of anatomical pathology, the historical a priori underpinning an alliance between seeing and saying is replaced by the primacy of the invisible over the sayable. Rather than the two-dimensional reading of a series of symptoms across a surface, the gaze of the anatomical pathologist maps a three-dimensional volume which includes invisible structures deep within the body (BC, xiii). “This structure, at once perceptual and epistemological, that commands clinical anatomy … is that of invisible visibility” (BC, 165). What Foucault is getting at is seen by considering a radiographer, whose gaze can, without interpretation or mediation, “see” hidden structures in images which, for untrained eyes, are only blurs. If the idea of the early clinic was of an “innocent eye”, carefully surveying its field, here Foucault is talking about a more active “educated eye” guided by what it knows, whilst not bringing that knowledge to consciousness (BC, 164). For Foucault neither perception nor conceptualisation has causal primacy. The object, mode of perception and concept are produced simultaneously, prior to theoretical elaboration (BC, 169).

Clinical practice strains at its conditions of possibility to find words for what might otherwise be imperceptible. It is not a matter of observing underlying essences beneath a disorderly appearance, or precisely matching sights with speech, but pushing the “foamy line of language, to make it encroach upon that sandy region” of clear perception for which words can’t easily be found (BC, 169). The clinician glances across visible surfaces to find words for the disease in the invisible causes now found in specific anatomical lesions. This further transformation was not “an act of psychological or epistemological purification” but rather the “syntactical reorganisation of disease in which the limits of the visible and invisible follow a new pattern.” (BC, 195). It is not a matter of observing more carefully or eliminating preconceptions. Rather, the historical a priori has shifted, giving primacy to the invisible over the sayable and extending concepts to carve the world differently. What appears as nothing more than fidelity to what is given to pure observation is actually a mode of perception based on a complex interpretive structure. By laying bare the a priori presuppositions of pure observation, Foucault’s work parallels the work of Kuhn, Feyerabend and Hanson on the theory-observation distinction (Kuhn 1996; Feyerabend 2010; Hanson 1958).
Foucault sees the fundamental structures of medicine changing with the development of a new style of thought and practice that placed emphasis on careful and precise observation. However, the belief in medical progress as simply better observation and abandonment of prejudices assumes “that the subject and object of knowledge remained what they were.” (BC, 137). Foucault doesn’t assume this. Rather than pure stable experience masked by shifting theories, both subject and object undergo a series of adjustments which bears on the objects given to experience, the classificatory grid and the subject’s position. 41 We see here Foucault spell out the conditions of possibility of knowledge in terms that exceed a simple opposition between subject and object.

One of the key innovations in pathological anatomy was the introduction of a distinction between phenomena associated with the disease itself and relatively autonomous processes leading to death. Death was no longer the limit beyond which there could be no medical knowledge but a legitimate object of inquiry, which gave a vantage point from which knowledge of life and disease could be obtained (BC, 145). The space of illness is now totally defined by the space within the body, not an external essence. “The idea of disease attacking life, must be replaced by the much denser notion of pathological life” (BC, 153). Only individual illnesses exist (BC, 168). This introduction of the individual into medical knowledge, by the thing-like presence of the dissected corpse, constituted man as an object of scientific knowledge, thus preparing the anthropological ground for the human sciences (BC, 197).

Modern medicine constitutes man as an object of science, as individuals subject to external norms of health. Out of the faded ideals of the Revolution grew values which expressed the positive significance of medicine. Rather than just curing ills, medicine “was given the splendid task of establishing in men’s lives the positive role of health, virtue, and happiness” by assuming “a normative posture, which authorizes it to dictate the standards for physical and moral relations of the individual and of the society in which he lives” (BC, 34). We see here themes that will come to the foreground in Foucault’s genealogical analyses. Disciplines governed by norms operate on individual bodies, constituting them as objects of human sciences which are both vehicles for, and effects of, power. Modern medicine embraces “a knowledge of the healthy man, that is, a study of non-sick man and a definition of the model man” (BC, 34). It adopts the normative stance of an arbiter of standards that defines positive norms of health and conceives life as the “bipolarity of the normal and the pathological” (BC, 35). As we will see, this distinction is fundamental to the human sciences of the 19th century, which are not only modelled on empirical sciences like biology, but on the positive

41 This theme anticipates Foucault’s later ethics, for example in The Government of the Living, in which the subject’s form, in its relation to truth, does not remain constant. (see Chapter 4)
norms of medicine (BC, 36). These norms also “linked medicine to the destiny of states” as “biopower” aimed at managing populations, in terms such as health and reproduction. Medicine is directed towards the “benefit of the state … the order of a nation, the vigour of its armies, the fertility of its people, and the patient advance of its labours” (BC, 35).

4 ORDER

OT, published in 1966, is an archaeological investigation into the human sciences, charting their emergence and placing them in relation to empirical sciences, philosophy and what Foucault calls the “counter-sciences”.42 Foucault proceeds with an analysis of the experience of “order” which constitutes the world.43 This order eludes consciousness of the scientist and yet is formative of scientific discourse.44 Against such constructivism, it might be objected that differences between things don’t rely on our categorical ordering, but are discovered to be already there as criteria are applied. Foucault would argue however, that whilst differences are real enough, it is humans who pick out particular differences, rather than others, on the basis of a particular ordering. Whilst the resultant categories are not arbitrary or unreal, neither are they entirely independent of us. By having a certain order we can make true statements about objects so ordered. However, the order could have been, if not for us being as we are, different. We can see this by looking at previous epistemes which order the world incommensurably differently. Foucault wants to bring this experience of the existence of order to his readers. It is an experience that implies an otherness which escapes the order imposed by scientific discourses. It “always plays a critical role” by opening up space to reveal the contingent basis on which knowledge becomes possible, sciences established and rationalities formed (OT, xxiii). By evoking an idea of otherness – a realm outside the discursive order of things - OT reiterates the Foucaultian themes of limits, transgression and freedom.

42 For Foucault, “empirical sciences” treat man as part of nature. His representations are products of an external world. They include biology, economics and philology. The “human sciences” are concerned with man as a subject, with his representations constituting the world. They include sociology, psychology and literary analysis. (OT, 384) By “counter-sciences” Foucault has in mind something like structuralist Lacanian psychoanalysis, Levi-Strauss’ ethnology and structuralist linguistics.

43 By “order” Foucault is referring to the underlying basis by which our concepts divide the world into objects about which we can make true and false statements. “There is no similitude and no distinction, even for the wholly untrained perception, that is not the result of a precise operation and of the application of a preliminary criterion.” (OT, xxi)

44 The idea of an underlying order strongly suggests Heidegger’s influence. (see The Age of the World Picture in Heidegger, 1977, 115-54).
OT charts how the Renaissance episteme gave way to the Classical episteme in the mid-17th century which in turn formed the background against which the modern episteme emerged. Each episteme is governed by an experience of order based on a different principle. Modern thought was not shaped through a progressive series of modifications and improvements, but completely transformed by the “mode of being of things, and of the order that divided them up before presenting them to the understanding” (OT, xxiv).

The Classical episteme embraced a principle of ordering based on representations of relations of identity and difference. Resemblances were analysed and differentiated by strict criteria. In principle this enabled complete certainty by the classification of representations in terms of differences mirroring the order of things in the world (OT, 61). Relations between things weren’t seen as just projections of a structure, but as resulting from their very nature (OT, 72). Prior to Kant there was no account of how consciousness comes to have a capacity to form representations. Consciousness or thought was necessarily representative. But towards the end of the 18th century, representation’s role as the unquestioned self-justifying starting point for knowledge became problematic and the modern scientific disciplines emerged in a total epistemological transformation (OT, 227).

The modern episteme orders things by analogies, such as functions between “organic structures” (OT, 236-7). Order no longer depends on the identity of elements, but the identity of invisible relations between elements within the object. What makes things stand out as objects is no longer necessarily represented. An object is no longer what it is because of its place within a pre-existing ideal system of classification, but because of its existence as a discrete structure given by its place in history. Rather than increased objectivity, more precise observation or more rigorous reasoning, it is History, the historical a priori of the modern episteme, that radically restructures thought to enable the emergence of new empirical sciences. Economics, biology and philology are not grounded on representations but something beyond – labour, organic structure, and the system of inflections.

As historical a priori, History plays a hidden role, constraining and enabling what can be said and thought (OT, 237). Its productive capacity enables, prior to all established chronology, the empirical domains to be what they are. It has both epistemological and ontological implications. It is both how we know things, “the space from which things come to knowledge” and “the mode of being of all that is given to us in experience” (OT, 237). Han suggests that this dual role can be understood by seeing Foucault’s position as analogical to, though not identical to, Kant’s understanding of causality. (Han

---

45 Lungs and gills, for example, are closely related by virtue of performing the same function, despite their markedly different properties (OT, 288).
Foucault is suggesting that History is the principle by which things are constituted precisely as the things that they are. We are bound to construe things as historical by the fact that we are governed by History. However, Foucault differs from Kant in that History, as historical a priori, comes to have and, may eventually lose, this status. The stability of Kant’s transcendental structures is purchased by the commitment to notions of universally shared cognitive faculties and the world-in-itself. Foucault is committed to neither notion. He insists that all we ever have is things as experienced by particular human beings in history.

Kant’s critique plays a pivotal role in the eclipse of the Classical episteme (OT, 263). He asks about the conditions for representation in general, showing us that thought is not necessarily grounded in representations, but something more fundamental, the synthetic activity of the thinker. By losing transparency and becoming problematic, representation becomes an object of inquiry, thought of as something firstly in itself, and then secondarily about something else. According to Foucault, this new order resulted in the fragmentation of the field of knowledge. A new distinction emerges between analytic and synthetic knowledge whereby maths and logic becomes sharply divided from the empirical sciences. The third form of enquiry is philosophical reflection which seeks a unified account of the nature of reality and the grounds of knowledge. Representation, no longer the unquestionable form of thought and knowledge, becomes a principal object of philosophical inquiry.

Together these three domains of enquiry – the analytic, the empirical and the philosophical - are the three dimensions comprising the modern episteme (OT, 378). Whilst they represent three irreducibly different forms of knowledge, Foucault notes that one can borrow methods from another or take another as its object of inquiry (OT, 375). The human sciences haven’t a place within modernity’s epistemological trihedron, but are insecurely located “in the interstices of these branches of knowledge … in the volume defined by their three dimensions” (OT, 379). The human sciences must be understood in terms of their special relations with all dimensions of modern knowledge. It is also in the human sciences that we can see the play of the aporia that emerged with Kant’s critique. I will now consider Foucault’s understanding of the modern empirical and human sciences, before turning to modern philosophy and its aporia.

---

46 For Kant, causality was neither in the mind alone, nor in nature alone, but was constituted as objectively in nature by the application of the pure categories of understanding.
Foucault understands modern sciences as decisive breaks with their classical counterparts. Rather than better observations or rational progress, what changed was “knowledge itself as an anterior and indivisible mode of being between the knowing subject and the object of knowledge” (OT, 274). Foucault is suggesting the autonomy and mutability of knowledge at the archaeological level seen by the emergence of the modern empirical sciences. In the following discussion, I will deal only with economics and biology and their related human sciences, since they provide a sufficient illustration of the pattern of emergences and the place of the sciences within the modern *episteme*.

**The Empirical Sciences**

According to Foucault, the Classical *episteme* analysed value solely in terms of exchange within a system of representation where money represented commodities. The decisive break was made by Ricardo who presented labour as not merely something exchangeable, but as the sole *source* of value, outside exchange. Commodities had value because people worked to produce them (OT, 277). What creates value is the accumulation of labour through the entire production process (OT, 258).

Economic history can now be understood as a linear causal series. Productivity of labour is related to “forms of production” – tools and machinery used, division of labour, capital invested etc. Since forms of production are themselves products of previous labour, current values are the result of a series of overlapping causes stretching back indefinitely (OT, 278). With man as an economic agent, economics can be understood historically. Man is not merely a bearer of representations, but is subject to the factors which caused those representations, such as the scarcity of nature or external threats. Ricardo views economic history as a history of increasing want as populations increase and resources diminish. Marx more optimistically heralds a new consciousness in which the arrangements men had previously attributed to nature are recognised as historically produced. Foucault provocatively suggests that Marx’s analysis is not the fundamental rupture with bourgeois thinking it is claimed to be (OT, 285). Ricardo and Marx share a view of economic life as the historical struggle of man to survive through his labour. Since both theories are founded on the epistemic break initiated by Ricardo, they fundamentally share the same mode of thought (OT, 280-6).

Similarly, modern biology emerges, with life no longer merely one category of natural things, but its specific object. For modern taxonomy, what is important is the functional similarities of organs, not identities and differences in phenomenal properties. Life in its non-perceptible purely functional form provides the basis for classification which depends, not on surface phenomena, but on elements hidden from view. Cuvier’s understanding of organic structures in terms of their functional roles, prior to, and independently of, taxonomy enabled a discontinuous classification of species to replace the
Species differentiation in modern biology is explained by the particular manner in which each species is linked to its environment, which is no longer merely a setting for the pre-determined essential natures of species. It views the essential nature of a species as *causally* dependent on its environment (OT, 298). As a consequence life becomes essentially historical, tied to *time*. Living things have been forced into discontinuous groups that have been formed in relation to the environment that existed at the time of their formation. Cuvier’s innovation set the stage for Darwin’s evolution.

Labour, living beings and languages are now understood as essentially historical realities, developing according to laws essential to them being what they are (OT, 319). The emergence of new empirical sciences involves a suite of new concepts, methods and objects (OT, 275). The new historical *a priori* – life, labour and language - were not objects waiting to be discovered, nor were they merely concepts. Rather, they are what Foucault calls “quasi-transcendentals” for contemporary thought (OT, 272). They function in the empirical sciences to provide the conditions of the possibility of the subject’s representational experience. Since the modern empirical sciences each have their own quasi-transcendentals, they lose the cohesion and possibilities of completeness and certainty they had in the Classical *episteme*. Knowledge becomes partial, tentative and disjointed.

**The Human Sciences**

Foucault sees the human sciences aspiring to, or borrowing from, all three dimensions of knowledge established when the homogeneity of Classical knowledge fragmented. However, the fundamental difficulties in defining the human sciences are the relations to the empirical sciences and philosophy (OT, 383). The human sciences borrow models and concepts which divide them into three interlocking “epistemological regions”, each corresponding to an empirical science. The psychological region is linked to biology and is concerned with man as a living being, in a capacity which “opens itself to the possibility of representation” (OT, 387-8). Psychology is not simply concerned with stimulus-response mechanisms like biology, but how such mechanisms are linked to representation. The sociological region is linked to economics, and is concerned with the way in which “the labouring, producing individual offers himself a representation of the society in which this activity occurs, of the groups and individuals among which it is divided, of the imperatives, sanctions, rites, festivities, and beliefs by which it is upheld or regulated …” (OT, 388). Sociology doesn’t seek law-like regularities in commodity transactions, but asks how things are valued and represented as commodities in the first place. Each empirical science contributes a pair of concepts in terms of which

---

47 For example, “vertebrates and invertebrates form absolutely isolated subareas, between which it is impossible to find intermediate forms providing a transition in either direction” (OT, 296).
the objects of the particular linked human science can be grasped. Psychology takes from biology the conceptual pair of functions regulated by norms. Sociology borrows from economics the conceptual pair of conflict governed by rules. Whilst the conceptual pair borrowed from each empirical science tends to dominate the human science linked to that empirical science, all conceptual pairs can operate in all human sciences.

Like philosophy, the human sciences are concerned with man as a subject and object, a knower whose representations constitute his world and enable him to know himself as an object within that world. However, the human sciences treat man’s representations as products and processes of unconscious structures. Meanings, functions and conflicts, organised by norms, rules and systems can all be represented without appearing in consciousness. Even though a particular society, individual or culture has no awareness of such representations, the human sciences can speak meaningfully of the function of a social practice, a conflict within an individual psyche or the meaning of a myth (OT, 394). By bringing to light functions, conflicts and meanings, human sciences develop an account of how man represents the fundamental realities of life, labour and language which appear in the empirical sciences as determinants of man as an object. These representations appear unconsciously as objects constituted by man as a subject who structures functions, conflicts, and meanings by the organising concepts of norms, rules, and systems. By the employment of higher-order organising concepts, the human sciences move towards the unconscious, and are able show how man represents the very forces that determine him as an empirical object. These concepts structure the entire field of the human sciences.

Foucault regards the human sciences as not strictly sciences, but belonging to a domain of knowledge by having a legitimate place in the modern episteme. The title “science” comes only from the models they borrow. Whilst they lack the predictive and explanatory capacity of empirical science, they are not mere opinions. They can yield objective knowledge by employing epistemologically sound methodologies. However, due to the interlocking of their interpretive models, the boundaries of the human sciences become blurred and “intermediary and composite disciplines multiply endlessly, and in the end their proper object may disappear entirely” (OT, 390). Methodological controversies frequently arise from disagreements about which constitutive model is most appropriate. The human sciences are intrinsically unstable, “treating as their object what is, in fact, their conditions of possibility. They are always animated by a sort of transcendental mobility. They never cease to exercise critical examination of themselves. They proceed from that which is given to representation

---

48 The models borrowed from empirical science are constitutive models. “They play the role of ‘categories’ in the area of knowledge particular to the human sciences” (OT, 389).
to that which renders representation possible, but which is still representation.” They are “constantly
demystifying themselves” (OT, 397).

The instability of the human sciences, with their blurred boundaries and proliferation of approaches,
does not derive, from the complexity and difficulty of their object, man. It is due to their location in
the modern *episteme* from which they are pulled in different directions by relationships they bear to
the empirical sciences, mathematics and philosophy (OT, 380). Instability does not prevent the human
sciences having a critical role in relation to the empirical sciences. They can lay “an invincible claim
to be the foundation of [the empirical sciences], which are ceaselessly obliged in turn to seek their
own foundation, the justification of their method, and the purification of their history, in the teeth of
‘psychologism’, ‘sociologism’ and ‘historicism’” (OT, 377). 49

Beyond the human sciences, Foucault identifies three broadly structuralist “counter-sciences” –
psychoanalysis, ethnology and a type of structural linguistics – as occupying a privileged epistemic
position. They possess “a perpetual principle of dissatisfaction, calling into question, criticising and
contesting what may seem well established” (OT, 407). They disrupt established disciplines and
settled beliefs across the range of human sciences. Foucault’s sketchy account of the counter-sciences
suggests that he is searching for a perspective that will enable a more radical critique of reason. I will
not discuss them further since their significance is little more than an undeveloped precursor to the
critical role developed in Foucault’s genealogy discussed in the following chapter.

**Man and the Analytic of Finitude**

I have discussed the modern *episteme* primarily in terms of the historical *a priori* of History which
Foucault introduces in Chapter VII of OT. But in Chapter IX he is concerned with another historical *a
priori* – Man, the object of the human sciences. By “Man” Foucault is not referring to humanistic
value concepts – a universal moral core, the dignity of the human person, the priority of a substantial
“I” or the freedom of a human subject. Nor is he referring to biological, social or psychological
notions. He is specifically referring to humans understood in a way not possible prior to the modern
*episteme*. He maintains Man is a “transcendental doublet”, the transcendental condition of all
knowledge of the world and, at the same time, a being in that world that can be known. He is
autonomous and rational, yet at the same time, the product of forces beyond his awareness. As such
“Man” does not designate a particular empirical object, but rather a question pursued by modern
philosophy which evades any stable solution.

---

49 This critical function can be seen in ongoing debates generated by the sociology of scientific knowledge. See
for example, Latour (1999).
Man’s emergence is directly linked to the decline of representation and its emergence as an object of inquiry. The representative condition of representations now appears to be outside representations, in the subject (OT, 259). What makes things thinkable isn’t part of what they are. Reality is represented only due to the contingent faculties of Man. Man is now a subject amongst objects, a subject not only seeking to understand the world of objects, but himself, in his very capacity as a subject-object. Man emerges with Kant as both the foundation of knowledge and an object of that knowledge. Man plays two roles. In order to be grasped empirically, science needs the transcendental arrangement of the human faculties (OT, 259). Whilst life, labour and language exist independently of Man as transcendental subject, they can only be known by the activities of that transcendental subject. This is Man in his foundational role. But Man is also an empirical object determined causally by life, labour and language within the field he himself has opened up. Man is not the object of any particular science but the surface effect of the superimposition of three empiricities – language, labour and life.

Biology, economics, and philology all show that Man is limited by the various processes by which he is enmeshed in the world. Science uncovers the laws that constitute man, charting the causal products of life, labour and language that go before Man and shape his representations (OT, 342). But rather than being trapped by this finitude, Kant turned it to an advantage, arguing that the very factors that limit our knowledge, that restrict it to the forms of space and time and the conceptual framework of the categories are, at the same time, the conditions for the possibility of knowledge (OT, 343). It was precisely these constraints that enabled objects to appear to us in the first place. What is specific to modernity is the doubling of Man, as both a transcendental subject and an empirical object of knowledge. This doubling is ambiguous since the distinction between the subject and object of knowledge, is grounded in the same being – Man. Man is employed to both separate and unite the empirical and the transcendental. The transcendental subject is the condition for knowledge of the empirical object, yet Man is one being.

From Kant onwards, philosophy has been employed to show on what grounds representations are possible and to what extent legitimate. Kant thought that it was precisely on the basis of universal and necessary limitations that we constitute objects, which could otherwise not appear to us. Whilst the empirical sciences of economics, biology and philology elucidate the laws by which Man is constrained and enabled by labour, life and language, so labour, life and language constrain and enable Man in knowing and clarifying those laws. The human sciences are specifically directed to, and take as their object, the constitution of Man’s subjectivity by these laws. Philosophy is also directed to the question of how Man constitutes the world of objects of which he is a part, although differently from the human sciences. Philosophy undertakes an “analytic of finitude”, an account of the relationship between Man as transcendental knower and Man as empirical object of that
knowledge. The problem is that the “positive” and the “fundamental” sides of Man must be viewed as the same (since Man is one being) yet different (since the fundamental grounds the positive). 50

One response by philosophers following Kant, such as Comte and Marx, has been the reduction of the transcendental to the empirical. The empirical conditions of the subject are put forward as the conditions of the subject’s knowledge. This reduction takes two forms: firstly, knowledge is explained in terms of the processes within the body which are involved in the production of knowledge (OT, 347). Secondly, knowledge is explained in terms of historical, social or economic conditions. However, both explanations uncritically assume that empirical knowledge of the body or society is just given, as though imprinted on a passive subject, then used to ground knowledge generally. Foucault refers to this sort of response to the analytic of finitude as “positivist”. The alternative to this is “eschatology” which claims the truth of scientific and historical accounts of empirical objects on the basis of the truth (once achieved) of our philosophical discourse about knowledge. 51

Another response is the cogito-unthought double which expresses the reality of man as both an experiencing subject, and the never fully understood object of that experience. The modern cogito cannot ensure epistemic immediacy and self-certainty because pre-reflexive conditions of knowledge obfuscate what appears as evident truths of reflection. Philosophers and human scientists seek clarity by subjecting obscure factual conditions to philosophical scrutiny, in a never-ending task of making present what is absent in the cogito, thinking the unthought. The problem is that gaining knowledge of the unthought within man is conditioned precisely by what remains unthought (OT, 352). The sciences of life, labour and language can reveal previously hidden ways in which our beliefs, desires, deeds and words are conditioned, but this knowledge is itself subject to further hidden conditions. It is not clear whether “I” is what I can consciously survey, or what is unknowable that always conditions, behind my back, such a survey. Man is formed by a complex network of background practices which he can never fully grasp, and yet he is the possibility of their elucidation. He is a product of a history whose beginning he cannot reach and at the same time he is the writer of that history. This tension between what we think we are and the unknown that conditions our thinking, forms a tight circle from which we cannot extricate ourselves.

A craving for explanation generates never-ending compulsions to keep uncovering what remains hidden in our nature. By incorporating this unthought in his knowledge, Man gains valuable insights

50 The term “fundamental” designates the actual existence of life, desire or language in its role as founding, or constituting knowledge, whilst “positive” designates this knowledge itself, as constituted by the transcendental activity of Man.
51 “Eschatology” is a lightly veiled reference to Marxism in which “the true discourse anticipates the truth whose nature and history it defines” (OT, 349)
into his actions. Science appears liberating by enabling modern reflective consciousness to adopt an objective stance on compulsions and motivations, suspending their force and making clear their causes (OT, 357). Under the banner of truth, thinking the unthought appears as a sort of political action promising eventual liberation as all that lies below our thoughts, deeds and words is brought to consciousness. Philosophical thought demands resolution of the question of Man by strategies such as positivistic or eschatological reductions or by attempts at making the unthought fully explicit.

Science provides the content of such strategies, though not as neutral descriptions. To think of human behaviours in terms of evolutionary psychology or neuro-physiology, neutralises their normative charge to mere events that happen. These are not neutral descriptions that leave things as they were but, as Foucault says, they “cannot help but liberate and enslave” (OT, 357). The problem is the attempt to make explicit something that, by its very nature, eludes reflection. Reflection doesn’t merely reproduce an original experience, but changes its given nature (OT, 357). Reflection is always already a theoretical attitude that involves an objectification of that which is its object. All expression, any attempt to put something in words is an objectification.

Beyond strategies of reduction, clarification, and interpretation employed by the analytic of finitude Foucault offers no further, more developed or fundamental solution to understanding Man. By the playing out of this analytic, philosophy falls into complacency (OT, 372). To move beyond this, we need to remove anthropology as a ruling category. Kant is clearly implicated in anthropology. Foucault notes that in his Logic, Kant added an ultimate question to the three critical philosophical questions which are “referred back to this fourth [question], and inscribed, as it were ‘to its account’: Was ist der Mensch” 52 Here Kant, on Foucault’s reading, places the question of the essential nature of man on the philosophical agenda of modernity as the most fundamental question. Kant’s question is still active, in constituting and supporting anthropology from which we are still not free (OT, 371).

The dangers become apparent when we consider that to speak on behalf of humanity is to invoke an ideological status of universality outside time. It is plausible to think that Foucault in OT, rejects Kant outright. This is certainly not the case with later Foucault. I will deal with questions of Foucault’s relationship to Kant in Chapter 4.

**The Death of Man**

Foucault claims that “man is an invention of recent date. And one perhaps nearing its end” (OT, 422). However, “these are not affirmations; they are at most questions to which it is not possible to reply;

52 In his introductory lectures on logic Kant writes: “The field of philosophy ... may be reduced to the following questions: 1. What can I know? 2. What ought I to do? 3. What may I hope? 4 What is Man? The first question is answered by Metaphysics, the second by Morals, the third by Religion, and the fourth by Anthropology. In reality, however, all these might be reckoned under anthropology, since the first three questions refer to the last.” (Kant, 1963)
they must be left in suspense, where they pose themselves, only with the knowledge that the possibility of posing them may well open the way to a future thought” (OT, 421). Foucault here is not predicting or describing what he thinks is happening. It is questions that “may well open the way to a future thought.” Foucault is urging this, for example, in passages where he contrasts our contemporary “twisted and warped forms of reflection” to “the unfolding of a space in which it is possible to think” (OT, 373). These are not neutral descriptions. Gutting argues that Foucault owes us reasons, which Gutting presumes to be found in non-discursive practices, bracketed in OT. (Gutting 1989, 224) Whilst the decentering of Man will change our understanding of knowledge, Gutting questions why Foucault urges change to a merely epistemological concept.

Whilst not spelt out in OT, the social and ethical implications of the concept of Man can be sensed in HM where we can see how human subjectivity becomes an object for a scientific gaze saturated with domination. This connection of knowledge and power will become thematic in Foucault’s genealogies, where the human sciences are linked to processes of normalisation. But even without cashing out a claim to liberation in social-ethical terms, surely Gutting would not object to the Death of Man thesis if he entertained Foucault’s claim that it would unfold “a space in which it is possible to think” (OT, 373). A further question is whether or not the “Death of Man” thesis implies the death of the subject more generally. Clearly there is much at stake. Concepts such as intentionality, agency, autonomy, responsibility and self-reflexivity are all linked to the idea of the subject and would be endangered. I argue that the Death of Man is a very specific claim that does not map onto the elimination of the concept of subject per se. Foucault is specifically rejecting Man, the simultaneously transcendental subject and empirical object of knowledge, as the historical a priori. Recognising archaeology’s subject as too passive, we will see Foucault later develop an account of a more robust subject capable of resistance and self-formation, a locus of freedom and responsibility, although such freedom and responsibility may always be found to be conditioned.

Whilst Foucault doesn’t endorse Kant’s dualism, nor does he reject it. He problematises it. In so doing, what he rejects is the hypostatisation of Man as a theoretical essence which, in the form of the analytic of finitude, requires one to resolve this riddle and find the truth as either a positivist or an idealist or as some other so far unimagined truth. Rather than “false”, I think Foucault would view the concept of Man as unproductive. It mires us in confusion and lends itself to the sorts of abuses we saw in HM. It traps us in endless circles, leaving no room to think critically. Foucault may well be an empiricist, but he doesn’t reject the transcendental. “In all of my work I strive instead to avoid any reference to this transcendental as a condition for the possibility of any knowledge. I try to historicise to the utmost to leave as little space as possible to the transcendental” (FL, 98). This isn’t a rejection of the transcendental but a methodological choice.
It is a pragmatic choice that is appropriate to his object – the present situation. In this situation the constraints from which Foucault wants to liberate us, arise from the hypostasised doubles of Man aspiring to provide an authoritative basis for the human sciences which lurch from positivism to idealism. The concept of Man is an obstacle to changing the present, “to imagine it otherwise than it is, and to transform it not by destroying it but by grasping it in what it is” (Foucault 1997, 108). Foucault however, is unable to say exactly what can and should be done or thought beyond the present since this is not a question he thinks we are able to address (OT, 272). So rather than offering a solution to the riddle, we must start from where we are, bringing to light the conditions of possibility for our thinking and acting and posing questions which “may well open the way to future thought” (OT, 421).

Foucault mentions one development which suggests the possibility of freedom - the re-emergence of a certain genre of avant-garde literature such as the works of Mallarme (OT, 327). Foucault wants to show not only how discursive limits of scientific knowledge and experience are constituted, but also, what escapes them. Whilst he shows how modes of subjectivity are constituted in scientific discourses, in OT he also gestures at how these limits might be transcended in avant-garde writing. In his genealogies, we will see literature’s role overtaken by an account of power and resistance.

Foucault doesn’t buy into the question of the “truth” of the subject, but is concerned with what is productive and life-enhancing. He is not putting forward any alternative theory of the nature of man. Any theory of man can potentially be stultifying. The one that stultifies our present is the one played out in the analytic of finitude. Rather than having an essence, human subjects are always constituted by, and embedded in, contingently evolving linguistic, historical, and cultural conditions. In a 1978 interview Foucault says “men are perpetually engaged in a process that, in constituting subjects, at the same time displaces man, deforms, transforms and transfigures him as subject. In speaking of the death of man in a confused, simplifying way, that is what I meant to say” (EW3, 276). Foucault’s Death of Man thesis is the recognition of this ongoing displacement, deformation, transformation and transfiguration, a recognition that is itself transformative. He doesn’t urge a “solution” to the analytic of finitude, but an overcoming by self-transformation which ultimately presupposes politics. Foucault’s proposed counter-sciences dissolve, rather than solve, the role of the concept Man by their external structuralist perspective. Rather than seeking something “deep” or “hidden” in human nature, Foucault calls Man into question, by exposing the conditions that make his reality as a representing subject possible, thus avoiding taking Man as a fundamental category (OT, 413-4). Foucault can only point to the exhaustion of the episteme and the future Nietzsche offers, as both a task and a promise, “the threshold beyond which contemporary thought can begin thinking again” (OT, 373). This future is articulated in Foucault’s later work as the continued undermining of essentialist thinking and a move towards a vision of relations of power that are sufficiently fluid for subjects to to reverse.
The emancipatory potential of Foucault’s archaeology is best seen through the perspective of his late work, particularly his idea of philosophy as a self-transformative exercise which I will discuss in Chapter 4. From this perspective, Foucault’s problematisation of the concept of Man aims to open a space for more illumination by exploration of the unthought in both theory and praxis. However, in posing the question of the demise of Man, it is not clear where Foucault sees himself located. Is his own thought inevitably determined by the epistemic order underlying his own archaeology? Is he located within the modern episteme, but with a vantage point from the cusp of a new episteme, from where he can see a new dawn lighting the horizon? It appears not, since he claims that epistemes form a barrier beyond which it is impossible to think. If we take this claim seriously, Foucault’s critique is as provincial as any other discourse, depriving him of any vantage to critique the human sciences and anthropology. Foucault later backs away from his claim about the pervasiveness of epistemes. We can also see that epistemes are not completely self-contained and all-embracing, since transitions across epistemological ruptures allow some continuity. Whilst Foucault compares discursive formations within different periods, he is not seeking complete pictures of these periods from a perspective outside all perspectives. By addressing a limited region of discourse, in order to reveal its contingency, Foucault isn’t necessarily undermined by the historical contingency of his own discourse. He is not a global relativist, bound to refute his own critique. It is always open to Foucault to claim the truth of his own critique which alerts us to the dangers of what we accept without question. The truth of his analyses and critiques is founded on the experience of his historically located situation, which he can acknowledge is as bound by its constraints as any other episteme. This doesn’t mean that it is relative to some pure state of affairs, untainted by any standpoint. I will return to the vexed question of Foucault’s relativism and foundations in Chapter 8.

Foucault’s insouciance might appear inconsistent with what some critics think is his ambition to provide a robust theoretical account of human beings, backed by philosophical justification. Rabinow and Dreyfus, for example, claim that Foucault’s archaeology is really attempting to provide a scientific theory, an attempt that fails, eventually leading him to abandon this project and take up

53 Rabinow and Dreyfus raise such concerns. See RD Ch.4
54 “In any culture and at any given moment, there is always only one epistememe that defines the conditions of possibility of all knowledge, whether expressed in a theory or silently invested in a practice” (OT, 183).
55 Four years later in the Foreword to the English edition of OT, Foucault cautions that the work is “a strictly ‘regional study’ and that such terms as “thought” or “Classical science” refer “practically always to the particular discipline under consideration” (OT x).
56 Adam Smith and J. B. Lamarck can both be considered as transitional figures between the Classical and Modern epistemes (see OT, 240-252).
genealogy (RD, Chapter 4). This criticism is misdirected. Foucault’s aim is not to develop scientific theory.\footnote{Foucault claims “my discourse ... is avoiding the ground on which it could find support” (AK, 226).} In fact, the scientific pretension of a theoretical understanding of human beings is precisely the target of Foucault’s critique. Whilst Foucault was drawn to the non-subjective external perspective that structuralism offered, there were very significant differences that he became increasingly frustrated trying to make clear.\footnote{See foreword to the English edition (OT xv).} He insists that the structuralist “counter-sciences” cannot be regarded as neutral bodies of scientific knowledge because they are tied to specific cultural practices - psychoanalysis to the doctor-patient relationship and ethnology to Western dominance at a certain historical moment. Psychoanalysis and ethnology both reveal a truth about the subject that is not an objective truth. Psychoanalysis for example reveals, as Maniglier puts it, an experimental truth in the sense that it only manifests itself in the actual transformations of the subject. (Maniglier 2013, 115)

Given Foucault’s later interest in practices of self-transformation, archaeology can be similarly regarded as a counter-scientific approach, critically oriented by its quasi-structuralist stance, and just as limited in its scientificity.

As I have argued, Foucault is best understood in relation to the French tradition of the philosophy and history of science. Archaeology dramatises the instability and contingency of knowledge which shifts, at the level of \textit{episteme}, in ways of which we remain unaware. Yet despite Foucault’s dramatisation of their contingency and instability, the human sciences remain sufficiently stable to ensure their authoritative power. This stability is conditioned by massive historical inertia which archaeology aims to unsettle. Archaeology develops a reflexive relationship to the contingencies that make us what we are, so that we can begin to transform those seemingly natural structures to which we are subjected. This approach is not aspiring to scientific objectivity or justification based on certain foundations, but attempting to come to terms with concrete historical data in a way that bears upon contemporary experiences of oppression and exclusion by transforming the present.

If we see the fact that archaeology doesn’t meet standards of scientificity or philosophical justification as a failing, this is perhaps because we are applying the wrong standard. Rather than asking what archaeology can justifiably know, we should ask what archaeology can do. By describing the ways in which scientific discourses produce subjects as their objects, archaeology can loosen the grip of those discourses. Although enlisting only one axis of the subject’s constitution (the axes of power and the self will have to wait for Foucault’s genealogies and ethics) archaeology is critically oriented towards the present, potentially transforming it by transforming us. This orientation does not entail the use of notions such as \textit{episteme} as totalising descriptions to explain why the past had to be as it was, or why the present must be as it is. Rather, such notions serve to escape the present illusion that we can’t think otherwise than we actually do. They are critical devices in a process which, by priming our
sense of contingency, destabilises tightly held epistemic commitments that appear as insurmountable obstacles. In Foucault’s histories of the present, we can only deal with problems as they appear to us as contingently located beings, with no absolute vantage point beyond this. From this perspective, no philosophical justification is required, or even possible.

OT’s critique of humanism and its philosophical underpinnings posits Man as something to be surpassed. Here is Foucault’s desire for freedom, to break free of whatever constrains us by virtue of our not seeing how it constrains us. In his archaeology, we see the beginnings of an approach which he was not able to clearly articulate at the time. Rather than the scientific approach to problem-solving that some of Foucault’s critics suggest, his project involves problematising what we take as universal and necessary. Foucault is not so much putting forward determinate claims, as opening up new areas for thought by a self-transformative process which experimentally challenges certainties. He is also gesturing towards an alternative conception of rationality. Foucault doesn’t see social scientific knowledge as the outcome of a progressive accumulation of individuals’ beliefs and activities tested against reality. Nor are scientific disciplines autonomously bound by their own internal logic. The scientific disciplines Foucault is concerned with constitute their own reality and interact with non-scientific discourses as part of a broader formation of knowledge (savoir) both propelled and constrained by anonymous unconscious forces, located in the rules of the episteme and its historical a priori which determine what can be said, seen and thought within a historical epoch. Whilst Foucault’s early archaeologies (HM and the first part of BC) included non-discursive practices and sometimes offer these as causes of change (though in strictly local and limited contexts), Foucault’s later archaeological histories assiduously avoid non-discursive practices. Here discursive practices float free, lacking any connection with non-discursive practices. By this bracketing, these works focus exclusively on the axis of knowledge. In his genealogy, he will fold archaeology’s knowledge axis into a vast web of non-discursive practices. Turning to Nietzsche, he will bring reason into relation with productive power, giving real force to the idea of the constitution of the subject.
CHAPTER 3
Science and Power

In the wave of unrest that followed May 1968, ongoing confrontations led to numerous arrests, and further protests against the denial of basic rights to those imprisoned. In February 1971, two months after his inaugural lecture at the College de France, Foucault founded Groupe d'Information sur les Prisons (GIP), aimed at giving prisoners a voice and disseminating information (Macey, 258). For the next 2 years, Foucault led an exhausting political existence of meetings and demonstrations. When in late 1972, prisoners started their own organisation, Foucault dissolved the GIP which had fulfilled its mission.

At his inaugural lecture, Foucault introduced the notion of a “will to truth” and linked discourse and knowledge to power (Foucault 1981). A conventional understanding suggests that science is exactly that form of knowledge that develops methods to abstract from the particularities of interests, prejudices and power in order to reach universal truths. Yet Foucault’s genealogy, developed in his lectures during the 1970’s and his two books Discipline and Punish (1975) (hereafter DP) and History of Sexuality vol.1 (1976) (hereafter HS1) argues that power and scientific knowledge are inextricably entwined. Genealogy chronicles how successive configurations of power and knowledge, tangled together in discourses, social institutions and practices interact to govern individual bodies and populations by constituting both objects and subjects. This chapter will examine the implications of Foucault’s account of scientific knowledge and its entanglement with power.

I will firstly discuss the continuity of Foucault’s project, how genealogy builds on archaeology. Nietzsche’s influence is clearly important. Next, I will turn to Foucault’s account of the emergence of modern power, and one of his most important innovations, the notion of power/knowledge. I will illustrate this by discussing the constitution and diffusion of psychiatric power/knowledge, a crucial component of the human sciences. I will then consider more specifically the constitution of subjects, objects, concepts and societies by scientific power/knowledge, before turning to two concepts that Foucault develops towards the end of the decade, biopower and governmentality, the latter serving as a conceptual hinge to his final phase of work. I will conclude with a discussion of some objections.

1 FROM ARCHEAOLOGY TO GENEALOGY

Archaeology and genealogy share an analysis of history emphasising the singularity of events. Rather than a complete innovation, genealogy is a return to, as well as an extension of, the approach of early
archaeological works such as HM. But now the simultaneous analyses of discursive and non-discursive practices becomes an explicit method. Neither power nor knowledge can be analysed alone, since they condition each other and are mutually generative. Neither is reducible to, nor explicable, in terms of the other. We have already seen that Foucault repudiates the distinction between science and ideology, a distinction that assumes the existence of a discourse governed by truth, free from power (PK, 102, 118; AK, 203-5). For Foucault, knowledge - not merely its applications, or choice of research questions, but the actual content of scientific knowledge – can’t exist outside power. By developing the concept of power/knowledge, Foucault can analyse knowledge and power as they transform and destabilise each other, so challenging the strict divide between the facts of power and the norms of objective knowledge. “Perhaps”, Foucault suggests, “we should abandon a whole tradition that allows us to imagine that knowledge can exist only where the power relations are suspended and that knowledge can develop only outside its injunctions, its demands and its interests” (DP, 27).

In OT, Foucault had restricted himself to describing systems of thought, explicitly bracketing explanation of changes from one episteme to another (OT English foreword, xiv ). Genealogy returns to what was bracketed, looking to concrete events and arrangements, to explain thought as not the product of thought alone, but as contingently developing from many small accidental causes, empirically accessible, though now forgotten. The conditions of possibility under investigation are no longer a set of epistemic rules, but a power/knowledge network consisting of all kinds of practices, “a heterogeneous ensemble of discourses, institutions, architectural forms, regulatory decisions, laws, administrative measures, scientific statements, philosophical, moral and philanthropic propositions, in short the said as much as the unsaid” (PK, 194). This dispositif, or apparatus, provides the conditions of possibility for the emergence of scientific objects and subjects. Foucault reveals this fine-grained multiplicity of factors which interact to cause changes beyond any subject’s intention or awareness. Rather than appealing to an ideal autonomous realm of thought or language, genealogy’s “gray, meticulous and patiently documentary” historiographical method holds metaphysics in check, by its analyses of the singularity of things, ideas, practices and experiences in terms of “events” (EW2, 369). As Paul Veyne puts it, events are like traffic accidents, unique and always due to the coincidence of a number of causal series (Veyne, 79). Rather than being subsumed under categories, meanings or historical essences, events undermine propositional knowledge whereby things fall within categories such that we are always dealing with the same thing. “Madness” is not always the same thing (BB, 3). What we now call “sexuality” doesn’t align with the “pleasures” of the Ancients, or the “flesh” of the Middle Ages (HS1-3). Foucault’s nominalism appears in a heightened register in his genealogies where, drawing inspiration from Nietzsche, he connects scientific discourses to mechanisms of power, explaining the origin of what we take to be natural in terms of a concatenation of contingent forgotten
events, such as the emergence of personages like the doctor, institutions like the asylum, spatial
arrangements, objects and concepts.

Foucault adopts Nietzsche’s claim that knowledge is an “invention”, so bringing it into relation with
instincts, interests and power (EW3, 6-8). Invention implies motives beyond knowledge for its own
sake. For Nietzsche, these were ulterior motives, animal drives to survive and dominate. Such motives
taint knowledge, stripping it of its claim to purity. Knowledge imposes order on a chaotic world so as
to dominate it (WK 203-4; EW3, 9). Against this picture, we want to insist that truth and knowledge
are precisely things not invented. No subject chooses what is true and false, like an inventor chooses
their materials. Foucault’s talk of “invention” challenges us to abandon the conventional perspective
of a subject of knowledge and adopt a more distanced view, perhaps that of an anthropologist. From
such perspectives, we can see that what appears so natural and given, is human, all too human.

But not just anything can be true. It is not particular truth claims that are invented, but the framework
of rules which distinguish truth from falsity. Within that framework, truth is not arbitrarily produced
(EW3, 230). And the framework itself is only arbitrary to a certain degree, since the truths which it
produces reflect a perspective constrained by the recalcitrant nature of things. There is a reality
external to the ordering framework, even if that reality cannot be wholly separated from the
framework through which we access it. Knowledge doesn’t mirror the world. Nor is it an instinct
(EW3, 8). Rather it is “produced because the instincts meet, fight one another, and at the end of their
battles finally reach a compromise” (EW3, 8). Inscribed at the level of instincts, knowledge comes
from struggle - compromised, provisional and temporary.

This theme of struggle is well illustrated in Foucault’s 1973-4 lecture series *Psychiatric Power*
(hereafter PP), which provides a Nietzschean account of the emergence of scientific knowledge from
fields of heterogeneous elements – various actors, institutions, objects – and its temporary
stabilisations and ongoing transformations. The struggles which constituted psychiatry came to a
climax with the play of power between patients and doctors, the initial basis of which was
psychiatry’s difficulty, in the late 18th century, in proving madness through demonstration (PP, 12,
133). An important breakthrough came around 1850 with the appearance of neurology which, it was
thought, could locate the mechanisms of madness in the pathological anatomy of general medicine.
Around the same time, the figure of the hysteric appears, offering clearly specified symptoms typical
of an organic disease, yet which constantly defy being fixed to an organic substratum (PP, 135).

A neurologist needs stable, coded, legible symptoms. Hysterical attacks must be regular and develop
according to a typical scenario sufficiently similar to a neurological illness but also sufficiently
distinct (PP p.340). With these conditions fulfilled, the psychiatrist can become a neurologist,
diagnosing a real illness. And the hysteric can claim the right to be properly ill, and so entitled to be free from asylum discipline (PP, 311, 316). But as this looks like collusion, a broader pathological framework is needed. Since there were no anatomical lesions to be found, Charcot looked for an event that could be assigned as a cause and developed the concept of “trauma” - a violent event, a blow, a fall, a fear, a spectacle – an event transcribed into the individual’s cortex, like an invisible pathological lesion which must be discovered in order to be certain that the hysteric is really a hysteric.

This requires that hysterics recount their lives, in order to bring to light this underlying event that persists, and continues to activate their symptoms, demonstrating that they suffer from an illness, best treated by neurologists (PP, 308). Here we see medical discourse elaborating a theoretical object, by a process made possible by the hospital structure and techniques of subjection. However, this discourse also generates a real object corresponding to its knowledge. But in a counter-move against this power of objectification, the hysterics poured out all the details of their sexual lives in words and gestures. To demonstrate that hysteria was a genuine illness, doctors needed an account free from excessive sexual elements. However, by this counter-maneuuvre the hysterics resisted objectification by giving the doctor more than he bargained for, and “finally [got] the better of the neurologists and silenced them” (PP, 322). Foucault’s account brings to light the interconnectedness of knowledge, power and resistance, how a shift in one precipitates a shift in the other in a struggle that is never finally settled. The struggle between doctors and patients “permeated the whole history of 19th century psychiatry” until it was finally conceded that Charcot produced the hysteria he described and that the neurological body eluded psychiatry (PP, 191-2).

The hysterics’ victory was countered by further tactical moves. The task was now not so much to eliminate psychiatric power which produced hysteria but apply it more precisely and effectively. The dominant approach, represented by psychoanalysis, sought to give neurotic symptoms full expression, but in a manner disconnected from the effects of power. It was thought that the power relation between patient and doctor could be sufficient to bring forth symptoms whilst not producing them. Thus we have the psychoanalytic conventions of free contract, the analyst’s silence, discursive freedom, etc, resulting not in an absence of power, but “a power that cannot be caught in any counter-effect, since it has withdrawn entirely into silence and invisibility” (PP, 343).

Like his archaeologies, Foucault’s genealogical histories are not moved by subjects, either individuals like Charcot or collective subjects like the hysterics. To this non-subjectivism, genealogy adds struggle and combat to destabilise foundations of knowledge by an account of historical beginnings “capable of undoing every infatuation” (EW, 372). Foucault’s ironic derisive critique opposes the
confidence of positivist histories of scientific progress, where truth is assumed to be an existing, though unknown state of affairs, towards which knowledge inexorably moves (WK. 208). Foucault suggests that ordering the world into conceptual categories establishes a form of knowledge that precedes truth, since truth relies on such ordering for its subsequent establishment (WK, 206, 227). By issuing from a particular conceptual order, knowledge necessarily “simplifies, passes over differences, lumps things together, without any justification in regard to truth. It follows that knowledge is always an overlooking”59 (EW3, 14).

Whilst power operates through concrete bodies unaware of its operation, Foucault doesn’t accept biological explanations of human behaviours (EW2, 380). His view of the malleability of what might be regarded as biological - sexuality, insanity or criminality - reflects Canguilhem’s idea that biological normality is a socially posited value (Canguilhem 1991, 160). Social and biological life are inextricably bound together. When certain traits come to be identified as “normal”, in the sense of what is common or average this fact can be adopted as a norm with prescriptive force which reinforces the trait. One only needs to consider the range of interactions between facts and norms surrounding notions of “fitness” to realise the social malleability of “biological” norms.

The body as the focus of the most minute scale of local practices is tied to large scale organisations of power (DP, 25). Both subjected and productive, it has become an essential component for the operation of a general mechanism of power and knowledge in modern society (DP, 26). The development of knowledge has not loosened the grip of power, but instead has become increasingly locked into its large scale organisation, potentially opening all domains to scientific investigation and control (EW2, 388).

2 THE EMERGENCE AND GROWTH OF POWER/KNOWLEDGE

Conventionally, “power” is seen as the exercise of intentional acts on others. Foucault’s analysis reveals power in many aspects that this conventional notion brackets out. Not a “theory”, but an “analytics”, it doesn’t ask what power is, but the “how” of “this grand, all-embracing and reifying term”. He asks how power flows, structures actions and moulds the constitution of scientific problems and their objects. Foucault charts the emergence of the dominant forms of modern power from earlier sovereign power (EW3, 336). Sovereign power is held by an individual or collective body and exerted repressively against others from a standpoint outside particular conflicts. It operates negatively by prohibiting, seizing things, bodies and life according to a binary system of permitted/prohibited (HS1,

59 I have adopted Kelly’s translation of “meconnaissance” as “overlooking” rather than the original translation’s “misconstruction” which does not fully capture Foucault’s meaning. (Kelly 2009, 167 n.20)
From the early 18th century, sovereign power declined in the face of complex social, political and economic relations that had developed. It has been complemented and surpassed by more efficient disciplinary power which operates automatically and invisibly (HS1, 88). Instead of the spectacle of the sovereign’s power, it is the subject that is made constantly visible, enabling prompt responses to any deviation by the application of continuous corrective pressures. Rather than “acquired, seized, or shared, something one holds or slips away”, modern power is applied to concrete bodies, mediated by alignments of a heterogeneous field of elements – actors, buildings, documents, tools, instruments, skills, practices and rituals - which are never completely stable (HS1, 94). It operates against a background assumption of other social actors who, acting in more or less predictable ways, unwittingly form part of a circuit, relaying anonymous power that has originated elsewhere. Power and knowledge are reproduced by the re-enactment of these alignments over time. Genealogy reveals the construction, maintenance and reproduction of these heterogeneous relations upon which power depends as never entirely stable.

The fine-grained “capillary interventions” of power must be distinguished from larger scale practices such as the law or class domination which are only the effects of a more pervasive power requiring more fundamental analyses (HS1 85, 95, PK 99). Foucault analyses scientific knowledge, inexorably entangled with power, by examining interactions between individuals at the local level, where a web of multiple force relations constitute a “micro-physics” of power (DP, 26). Since power emerges from multiple interactions - supporting, cancelling, amplifying, undermining, relaying, extending etc - it cannot be analysed or reduced to a binary opposition between rulers and ruled (HS1, 92).

Disciplinary power is disseminated throughout society to not only constrain but produce and enhance knowledge, health, wealth and social systems. The individual is an effect of power, not raw material pre-existing the power that moulds him (DP, 170). Discipline operates by a “subtle coercion” of the individual body – its movements, attitudes, gestures - as an object to be manipulated. Disciplinary power is meticulous, exhaustive, all-pervasive and continuously exercised, constituting bodies as docile by resolving resistances, and useful in terms of skills and aptitudes (DP, 138). Bodies are separated and aggregated both conceptually through classification, and concretely by spatial-temporal distributions. These subjects are constituted through discursive objectification and bodily manipulation oriented towards optimal states (DP, 141- 162). Material subjection enables theoretical objectification and the emergence of human sciences like criminology, pedagogy and psychiatry. These sciences further facilitate the refinement of disciplinary technologies which mark bodies with qualities that make them distinct individuals who appear in a grid of relations, functions and ranks in a

---

60 For example, an employer writing a reference is not simply the holder of power exercised on the employee. Her action depends on the maintenance of the institution of referencing and the ongoing recognition of the authority of her position. These are clearly unstable social alignments which can be reversed.
networked segmented space. By this “assignment of subjective positions”, individuals are allotted roles in the social world, providing a range of possibilities comprising specific arenas for the subjection to, and exercise of, power (AK, 107).

In DP, Foucault treats humans as passive, objects to be moulded by disciplinary power which operates as a continuous uninterrupted supervising coercion (DP, 137). However, we shouldn’t regard this treatment as Foucault’s view of the truth of the human condition without remainder. Certainly this bleak picture of human subjectivity dominates DP. But we must also recall Foucault’s lectures at the College de France in the years immediately preceding the publication of DP which lay out the struggles between established institutional powers and the resistances of the marginalised. As we have seen with the hysterics, in some instances resistance succeeds, at least temporarily. Immediately after finishing DP, Foucault set to work on the History of Sexuality vol 1, in which again we see that “power is never external to resistance” (HS1, 95). Foucault’s subsequent acknowledgements of DP’s shortcomings suggest to me more recognition of an over-emphasis on the passivity of the subject than the acknowledgement of a previous view as false. “When I was studying asylums, prisons, and so on, I insisted too much on the techniques of domination” (my emphasis, cited in McCarthy 1992, 279 n.64) As we will see by “domination”, he meant asymmetrically frozen hegemonic power relations, not relations of force per se.

Foucault attributes the success of discipline to three interacting instruments - hierarchical observation, normalising judgement and the examination. The Panopticon, Foucault’s well-known model of hierarchical observation, illustrates the economy of disciplinary power (DP, 202-5). Power is directed towards those unavoidably visible objects, for who this power of objectification is invisible. This invisibility arises not from ideology or false beliefs, but the spatial configuration of power and its imperative of efficiency (DP, 187). The observational gaze operates as a network in which every individual is at the same time both object and subject of power (DP, 176-7). The prisoner is watched, but watches himself. 61 The warder is the watcher and yet he too is watched. Whilst the prison governor is at the apex of the prison’s hierarchical structure “it is the apparatus as a whole that produces power and distributes individuals in this permanent and continuous field” (DP, 177). On one hand, power “is everywhere and always alert” and on the other hand, is “absolutely discrete, for it functions … in silence.” By its omnipresence power becomes invisible.

61 “There are two meanings of the word ‘subject’: subject to someone else by control and dependence, and subject tied to his own identity by a conscience or self-knowledge.” (EW3, 331)
Judgement directs observation to what is relevant, just as what is observed forms the content of judgment. Judgement “measures in quantitative terms and hierarchises in terms of value the abilities, the level, the ‘nature’ of individuals.” This judgement measures incremental deviations from a norm, automatically introducing “the constraint of a conformity that must be achieved” (DP, 183). The slightest deviation from the norm is corrected (DP, 177-84). Whilst normalisation imposes homogeneity, it individualises by introducing all the shading of individual differences (DP, 184). The techniques of observation and judgement combine to form the examination - an exercise of power over individuals and an experiment to find their truth. The examination gave rise to the human sciences (EW3, 59). It came to take the forms of “tests, interviews, interrogations and consultations” that reproduced and extended power/knowledge relations throughout all disciplinary institutions by tying power of subjection to the knowledge of individuals as objects (DP, 184-5, 226-7).

Whilst disciplinary power arose in isolated institutional islands – prisons, asylums, schools, barracks, hospitals - it was disseminated by “a sort of general parasitic interference with society … a disciplinary society replacing a society of sovereignty” (PP, 66). Although many practices of sovereign power remained in place, they now function by the social cohesion provided by disciplines operating at a different location and scale. Foucault’s 1974-5 lecture series Abnormal (hereafter AB) illustrates this dissemination by focusing on the diffusion of psychiatry beyond the asylum. Without resting on any single institution, psychiatric normalisation extended itself by interactions between institutions – church, state, schools, prisons, law and medicine. Psychiatry underwent mutations and spawned other sciences. These normalising disciplines came to permeate the entire social body (PP, 86). Today the conceptual and normative structures embodied in these disciplines permeate modern cultures. Psychologistic and quasi-Freudian categories constitute categories of subjects and objects to form the barely recognised background of interpretations of selves and others.

Foucault explores this diffusion by the genealogies of the figures of the human monster and the little masturbator, produced by particular configurations of power/knowledge. The principle of the monster’s transformation is found in shifts in the power to punish, which from the 19th century, was not applied merely according to the act, but increasingly according to the criminal’s nature. This new disciplinary logic required the scientific study of man and the stipulation of the modern legal subject – a universal subject endowed with the capacity for autonomy and freedom, but also the capacity to be trained and corrected. Man both discovers his nature and manipulates it to achieve his potential. By recourse to this universal subject, the development of scientific knowledge about man became possible. Legal procedures – witness accounts, investigations, examinations - produced the truth regime of the sciences represented in a courtroom by sanctioning types of truthful discourse to justify penal practices. The interaction of police, judiciary and psychiatrists determined the rules for the formulation of valid propositions which generated the discipline of psychiatry.
To be constituted as a proper science, concerned with the protection of society, psychiatry must show that it can detect danger not apparent to others. Thus drawn to the problem of criminal madness, psychiatrists argued that the motiveless crime was impelled by an irresistible inner force – instinct - that could break through all barriers of morality and reason, even the desire to live (AB, 119-21). The concept of instinct becomes the vector of abnormality, on the basis of which psychiatry not only brings serious disorders, but minor irregularities, within its ambit as mental illness (AB, 132).

This expansion is assisted by a field of new symptoms, consisting of deviations from a range of norms situated on the axis of voluntary/involuntary. With all conduct located on this axis, ultimately “there is nothing in human conduct that cannot be questioned by psychiatry” (AB, 160). Psychiatry could be linked to organic and functional medicine by the understanding of neurological disturbances of voluntary conduct. This means that psychiatry was able to exploit the norm in the sense of a principle of an appropriate and adjusted functioning. The explosion of symptoms also covers the growing array of conducts made possible by the norm, understood as a rule of conduct or principle of conformity. These two uses of “norm” are conflated, mutually adapted and partially superimposed (AB, 162). Thoroughly underpinned by this interplay, psychiatry treated the most common everyday conducts as both deviations from a rule of conduct and pathological dysfunctions. Supported by this play of meanings, psychiatry becomes medico-judicial in its nature and everyday practice.

The lineage of Foucault’s other figure, the little masturbator, can be traced back to medicine’s taking control of sexuality from the Church. From the mid-18th century, a great anxiety had arisen about child masturbation which, like sexuality generally, was increasingly seen to be within the province of medicine. By the late 19th century, sexuality is seen as the root cause of all forms of abnormality and, by its intrusion into the family, psychiatry takes up the figure of the little masturbator. The two figures – the human monster and the little masturbator – now merge to become the modern character of the abnormal individual. Psychiatry is able to expand, to intervene in both the family and the judicial system by unifying instinct and sexuality. It develops discourses, methods, analyses, concepts and theories which combine sexual instinct not only with illness but with an entire spectrum of abnormality from minor irregularities within the family to monstrous crimes (AB, 277). Here we see modern scientific power/knowledge, operating autonomously and independently of the intentions of actors, embedding, maintaining and expanding itself in an ongoing feedback loop. “Modern power fashions, observes, knows and multiplies itself on the basis of its own effects” (AB, 48).

Foucault argues that the understanding of the human sciences cannot be divorced from their association with disciplinary technologies. Psychology, psychiatry and criminology were originally located within disciplinary institutions such as hospitals, prisons or asylums, which needed more
effective discourses, knowledge and practices. Such disciplines developed as specialisations with their own methods and standards of evidence, not simply as the progress of reasoned discovery, but in close association with disciplinary technologies. Unlike the natural sciences, Foucault thinks that the human sciences have not separated themselves from their emergence within disciplinary technologies directed towards producing docile, useful bodies (DP, 305).

The continuous expansion of the human sciences was enabled by clinical sciences of the individual which provided a vast compilation of documentation, further refining disciplinary techniques. By capturing individuality in documentation, the examination homogenises individual features into codes. Elements can now be correlated and organised “into comparative fields making it possible to classify, form categories, determine averages and fix norms” (DP, 190). The examination constitutes the individual as describable and analysable, a constant object of a permanent body of knowledge, “a reality fabricated by this specific technology of power” constituting individuals as “cases” which are objects of both knowledge and power (DP, 194). Anticipating his concept of biopower, Foucault discusses how documentation also constitutes populations by comparative systems that measure “overall phenomena, the description of groups, the characterisation of collective facts, the calculation of gaps between individuals, the distribution in a given ‘population’” (DP, 190).

Whilst Foucault traces the human sciences to the disciplinary technologies of the 19th century, the natural sciences arose from an earlier technology of power, the inquiry - “another power, another knowledge” (DP 226). The inquiry had been a particular practice of judicial power, which emerged in the shift to juridical law requiring the establishment of the objective facts of “exactly who did what, under what conditions, and at what moment” (EW3, 5). Whilst the examination was organised around the norm, the inquiry was directed at what happened by the evidence of those who should know. It spread into areas not directly connected to the exercise of power, migrating into the sciences of observation - geography, botany, zoology, astronomy or medical sciences (EW3, 4; DP, 225).

Foucault refers to the increasingly fine grid of scientific techniques of the inquiry which ultimately leads to “a police kind of investigation into people’s behaviour … [revealing] the tightening grip of the technique of inquiry in our kind of society” (PP, 246). It seems that here Foucault is referring to human sciences that model themselves on the empirical sciences and claim a similar authoritative validity. This colonisation extends both in depth, in the degree to which it penetrates bodies, actions and thoughts of individuals, and in breadth as “a planetary extension to the entire surface of the globe” (PP, 246). The expansion of techniques of inquiry seems to have no limit (PP, 246). Whilst this scientific truth of inquiry is in principle universal and accessible to everyone, it is also rare, buried, and difficult to reach (PP, 247). Universities, learned societies, schools and laboratories are “all ways of organising the rarity of those who can have access to a truth that science posits as universal.”
Individuals only have an abstract right to be a universal subject, since concretely, such organisations “will necessarily entail rare individuals to perform the function of universal subject” (PP, 247). In addition to highlighting the power relations that underpin the structures of elite knowledge, Foucault is raising concerns that we will see in Habermas in relation to the increasing specialisation and remoteness of science from the lifeworld.

By charting a separate genealogy of the natural sciences, Foucault is not positing a rigid division between the human and natural sciences. As we have seen in chapter 2, the human sciences are located precariously between the three branches of the “epistemological trihedron” constituted by philosophy, mathematics and the natural sciences (OT, 378). The human sciences rely on the natural sciences for their models and guiding concepts. As we will see, at the broadest level of analysis, this interactive web includes technologies, concepts, objects, which lock together to constitute modern society as “the carceral archipelago” (DP 297).

3 POWER/KNOWLEDGE AND THE INTELLECTUAL

Thus far, we have discussed Foucault’s notion of power/knowledge with little regard for the fact that it cuts across our conventional understandings of science as explicitly distancing itself from power. In Foucault’s genealogies, we can see how shifts in scientific knowledge are linked to shifts in power. We saw how the inquiry gave rise to the empirical sciences, whilst the examination accompanied the human sciences. Surveillance, normalising judgement, examination and documentation are all both exercises of power and knowledge which have no separate concrete existence. Force relations are not outside, but immanent to, not only knowledge but relationships such as economic or sexual relationships (HS1, 94; STP, 2). However, these relations don’t reduce to power and power cannot fully explain them.

Along with new forms of social control, new sciences emerge with new objects, such as the homosexual and delinquent (DP, 138). From the 18th century, a threshold was crossed after which “the formation of knowledge and the increase of power regularly reinforce one another in a circular process” (DP, 224). By their tightened relations, power and scientific knowledge were now mutually constituted. More extensive finer-grained knowledge enabled more continuous and pervasive exercise of control in a sort of political-epistemic homeostasis around a norm. Documentation becomes a resource for ongoing and more precise examinations, enabling the emergence of new discourses about inner motivations, personalities, predispositions and desires, leading to an expanded and more intense subjective life. Individuals began to consider themselves as objects for attentive observation. Human
sciences produce not only theoretical objects of science, but also of reality – desires, forms of experience, kinds of bodies.

At the conclusion of this chapter, I will consider the concern that Foucault’s notion of power/knowledge undermines his entire project by denying the epistemically normative stance from which he can argue. For now, we can see genealogy as normative, in the sense of being concerned with showing us how we can know better. Foucault’s genealogy is aimed at questioning what the human sciences have simply assumed, for example, the ahistorical nature of desire, or the objective nature of sexuality. By reflecting on how a belief about an object is related to power, we can be better justified in granting a truth value to the belief or an ontological status to the object. More often, genealogy enables a revaluation of an object. Genealogy is somewhat similar to Karl Mannheim’s “unmasking turn of mind” which does not seek to refute concepts as fictitious, but rather displays their “extra-theoretical function” so stripping them of their false authority.

Foucault’s critique doesn’t necessarily show the human sciences to be untrue or even dubious. Genealogy problematises the human sciences, stripping them of their automatically granted authority, and opening them to a broader contextualisation. It shows how power/knowledge constructs objects which, endorsed by scientific authority, become naturalised - objects like sexuality, delinquency, madness, the economy and the market. Such objects don’t exist in an objective scientific sense. They are not independent of our recognition and naming them. But once recognised and named, they are real enough. When knowledge is granted the honorific title of science, it becomes a foundation on which we build further knowledge and practices. Foucault’s genealogy problematises by showing how these foundations were constructed. Genealogies operate as a counterpoint to the familiar stories about liberation from sexual repression or the enlightened humanisation of punishment. Describing practices which further entrap us in repressive regimes of discipline, Foucault’s accounts seem to call for resistance. But he rejects the idea that there is any position outside power. Power is always accompanied by resistance which is never exterior to power (HS1, 95). Resisting power is simply subjection to another power. “Power is ‘always already there’ [and] one is never ‘outside’ it.”

---

62 For example, aspects of sexuality, such as gender orientation, have frequently been revalued.
63 For a discussion of Mannheim in this context see Hacking (1999, 53-6).
64 Problematisations are “the way an unproblematic field of experience, or a set of practices, which were accepted without question, which were familiar and ‘silent’, out of discussion, becomes a problem, raises discussion and debate, incites new reactions, and induces a crisis in previously silent behaviour, habits, practices, and institutions” (Foucault 2001, 74).
But importantly, this “does not mean that one is trapped or condemned to defeat no matter what” (PK, 141-2). In his later work we will see Foucault distinguish between domination “which people ordinarily call power” in which asymmetrical power relations remain frozen, and power relations “understood as strategic games between liberties – in which some try to control the conduct of others, who in turn try to avoid allowing their conduct to be controlled or try to control the conduct of others” (FL 447). Rather than a conflict between opposing forces, the exercise of power is now “a mode of action on the action of others.” And “action” implies freedom (EW3, 342). Thus power never exhaustively determines a subject’s possibilities. So whilst power is everywhere, even in the most harmonious and natural relations, this isn’t, like Weber’s iron cage, a situation we are trapped in.

Given the omnipresence of power in the broader sense of force-relations (which includes resistance) how does Foucault, how does Foucault understand the intellectual’s role, his own role, in relation to the present? Not only is knowledge an outcome of battle, but it is a weapon in the battle, and genealogy is no exception. In the 1975-6 course Society Must be Defended (hereafter SD), Foucault opposes his “historicism” to the anti-historicism of “all the great philosophies” and the human sciences which adopt the pose of the universal man beyond history. Against the tyranny of these “overall discourses”, Foucault urges an “insurrection” of “subjugated knowledges … of the psychiatrised, the patient, the nurse, the doctor, … the delinquent” (SD p.7). His genealogies are not merely descriptions of power/knowledge from an external perspective but “weapons in the struggle … we wage war through history” (SD, 172). They impel action by problematising the power monopoly of established sciences by playing off disqualified knowledges against unitary scientific theories which “filter them, organise them into a hierarchy, organise them in the name of a true body of knowledge, in the name of the rights of a science that is in the hands of the few” (SD, 9). “Global theories” that seek hegemony by curtailing, dividing, caricaturing and overthrowing non-global discourses are a “hindrance to research” (PK, 81). Distortions and omissions are required to contain particular concrete things and events within the terms of a unitary science. By resuscitating historical knowledges submerged in totalising schemas, genealogy does not undermine the content of science, or relativise its truth claims. It directs its insurrection against “the centralising power-effects that are bound up with the institutionalisation and workings of any scientific discourse organised in a society such as ours.” Genealogies question “the aspiration to power that is inherent in the claim to being a science” (SD, 10).

Foucault’s problematisations challenge not so much the universalist aspirations of science, but the accompanying sense of ahistoricity and necessity which serve to legitimate the application of unitary

---

65 By “subjugated knowledges”, Foucault means historical contents that have been disqualified as below the level of scientificity, or masked by formal disciplines which subsume them within universal schemas (SD, 7).
sciences as blunt instruments of social engineering. Instead he advocates responding to local issues with a specificity that escapes global systems of thought. The “specific intellectual” works “at the precise points where their own conditions of life or work situate them (housing, the hospital, the asylum, the laboratory, the university, family and sexual relations)” (PK, 126). “It is not a matter of a battle ‘on behalf’ of the truth, but a battle about the status of truth and the economic and political role it plays” (PK, 132). He is not so much concerned with what is true or false, as the social and political functions of what is taken as true or false. Foucault wants to draw our attention to the pre-conditions which social, economic and cultural power establishes for what can count as true, not to free knowledge from power altogether (since this is impossible), but to loosen the grip of the current hegemony to reveal the possibility of thinking beyond the present.

Whilst Foucault was critical of the “universal intellectual”, he wasn’t just a specific intellectual who, on the coalface of struggles, galvanises support for the marginalised by revealing their forgotten histories. To some extent, Foucault’s project was that of a specific intellectual. But in its generality it was more a meta-critique of traditional humanist critiques. Rather than a local critique, it offered a panoramic account of the operation of power/knowledge, not only in prisons or asylums, but generally throughout modern society. It is in this sense that he refers to his books as “a kind of toolbox others can rummage through to find a tool they can use however they wish in their own area.”

4 THE CONSTITUTION OF SUBJECTS, CONCEPTS AND SOCIETY

The subject – the delinquent and women refugees

Thus far we have considered Foucault’s account of modern power/knowledge, along with genealogies charting the emergence of specific sciences from an agonistic field of shifting relations between actors, objects, concepts and institutions and their dissemination by technologies of power. I will now focus in more detail on how power/knowledge constitutes subjects, before turning to the constitution of concepts and societies as wholes. In DP, Foucault offers a “history of the modern soul and a new power to judge” deriving “its bases, justifications, and rules” from the “scientifical-legal complex” (DP, 23). By “soul”, Foucault is not suggesting a metaphysical illusion, but something very real, a historical reality corresponding to “a certain technology of power over the body” (DP, 29). The modern soul is produced by the functioning of power/knowledge that observes, supervises, arranges, judges, and corrects. It is an effect of a certain type of power and the reference of scientific

knowledge. But whilst this knowledge is not a mere illusion, neither is it the discovery of a “real” nature of man’s soul.

The term “soul” usually connotes that which is most authentic, revealed in the emancipation from the distorting influences of power. But Foucault insists that this soul is already constituted by power, outside of which, there is no subject. He doesn’t think that human progress consists in the gradual emancipation from power of natural qualities of rationality or autonomy. The modern subject is a historical achievement, produced through power/knowledge that examines and judges it against norms which introduce “all the shading of individual differences” (DP, 184). The modern individual is the effect and object of an overlapping of power and knowledge, strategies of discipline and human sciences.

The delinquent is one such effect of power and object of knowledge. Foucault argues that around the time the prison emerged, illegalities had to do with conflicts between social classes. By constituting the concept of a criminal underclass, comprising easily identified and controlled “delinquents”, these conflicts could be resolved, though not in the sense of being eliminated. According to Foucault, science doesn’t discover delinquency which the prison then addresses. Prison, supported by human sciences, creates delinquents, picking out and isolating them physically and conceptually as the "abnormal" elements of society. In a field of concepts, institutions and objects, delinquents are constituted as something to be managed. The prison persists because “in fabricating delinquency, it gave to criminal justice a unitary field of objects, authenticated by the ‘sciences’, thus enabling it to function on a general horizon of truth” (DP, 256).

One way to grasp this notion of “fabricating” subjects is by considering what Hacking calls “interactive kinds” (Hacking 1999, 103-9). Unlike natural kinds, the concepts of which can be applied to situations prior to the emergence of those concepts, social objects seem to emerge along with their concepts. Hacking, as an example, argues that whilst women have long been fleeing their homelands, it may not be right to call them “women refugees” because this concept has only recently emerged from a social matrix comprising institutions, actions of advocates, judges, journalists, politicians, concepts such as rights and objects such as borders. A woman fleeing her homeland is not a refugee until she falls within the appropriate categories, and this involves appropriate interactions with objects and agents. Subjects of such categorisation may, perhaps unconsciously, conform to the expected stereotype, or disguise or exaggerate actions, like Charcot’s hysterics. Categorisation derived from scientific knowledge is applied to bodies, moulding them to norms and changing the way people experience themselves. Changes in behaviour may then require a refinement of the original categories. This is what Hacking calls the “looping effects” of interactive kinds, the kinds that the human sciences take as their objects.
It is in this vein that Foucault thinks we need to discover how “subjects are gradually, progressively, really, and materially constituted through a multiplicity of organisms, forces, energies, materials, desires, thoughts etc” (PK 97). He claims it was only in the 19th century that “the homosexual [became] a species” (HS1, 43). Before this, power/knowledge had not aligned the contingent configuration of “organisms, forces, energies, materials, desires, thoughts etc” for the homosexual to exist as today. Of course it is not only the homosexual or delinquent that has emerged. As Hacking argues, who we are is not just what we do or did, but also depends on the possibilities of what we could have done. If it wasn’t possible to be a delinquent or homosexual before the 19th century, it also wasn’t possible to be a non-delinquent or heterosexual.

Such unregistered shifts in what we are, are generated by the increase and spread of discipline which reinforces itself by structuring situations as occasions for more discipline. A well disciplined individual incorporates not only specific capacities of training, but discipline as such. The population submits to normalisation by various disciplinary techniques within schools, factories or offices. However, we are not trapped within a power regime, since discourses are “tactically polyvalent”, able to be used for different purposes, including resistance (HS1, 100). A “reverse” discourse revalues by recasting normalcy, insisting that what has been ruled abnormal is normal. Canguilhem saw the norm as unstable. “The possibility that the norm offers to regulate leaves room for another possibility, its opposite” (Canguilhem 1991, 240). Similarly, Foucault viewed many biological norms as representing not an objective standard but a desired state of affairs, irreducible to objective scientific concepts (Canguilhem 1991). Whilst normative judgements on health are “subjective”, they are not mere whims. A doctor may correctly tell a patient that he is sick based on physiological results, even when he feels healthy. But this is only because science has uncovered causal connections between the current “healthy” physiological state and one that any patient would call “unhealthy”. This apparently objective diagnosis results from uniformity in judgements about organic illnesses. Everyone feels the pain of appendicitis. The objective appearance of such judgements allows their slide into judgements on which there need be no uniformity, for example, judgements on gender orientation. Here we see the norm, derived from what is frequent, become valued as a standard which must be enforced and maintained under certain historical and social conditions. Yet gender orientation is more deeply subjective, more part of the essential identity of an individual than a purely organic condition, such as appendicitis, which may be regarded as an accident. Foucault accounts for our most intimate and personal experiences as conditioned by large-scale historical events of which we are only dimly aware. He is concerned with the conditioning of these subjective identities by normalisation which directly on what an individual appears to be “by nature”, an appearance which obscures the “second nature” of normalisation that has produced her.
The Concepts of Sexuality and Sex

Scientific discourses constitute not only subjects and objects, but concepts such as “sexuality” and “sex”. In HS1, Foucault maintains that these are not what the human sciences assume them to be. His point of departure is the so-called “repressive hypothesis”, the idea that beginning in the 18th century, sexuality has been opposed, silenced, restricted and regulated, a situation we must liberate ourselves from. Contrary to this, Foucault thinks that there has been an explosion of discourse on sex since the 17th century, beginning with the counter-reformation’s prescriptions about confession. Practices of self-examination applied increasingly finer distinctions, leading to a self, not discovered but constructed by the categories applied. Instead of repression, there is an increasing incitement to discourse (HS1, 34). This discourse posited sex as a powerful and irrational force that required techniques of confession to control. Confession became generalised throughout modern society, particularly in psychiatric and psychological disciplines. What started as the ritual of confession under church power eventually became constituted in scientific terms.

Sexuality, based on private sensations of pleasure, secret fantasies and dangerous excesses, developed from the proliferation of scientific discourses, produced and harnessed by power from the beginning of the 18th century. As an interface between concepts, sexuality constitutes a network that interconnects various styles and perversions into a single grid of scientific power/knowledge (HS1, 44, 103). On Foucault’s account, any attempt to liberate our “true” sexuality, is to simply fall under the sway of another strategy of power. Mechanisms of power produce, through scientific discourses, the ways in which we experience and conceive sexuality. Any form of sexuality, “repressed” or “liberated”, involves internalising categories and norms provided by power/knowledge. The belief that one is resisting repression supports domination by concealing the workings of power.

Foucault considers an objection that focuses on the ontological status of sex: that his history of what is said about sex is irrelevant to the biological reality of sex itself (HS1, 150). By ignoring this reality, it is objected that Foucault hasn’t come to grips with the real referent of discourses he analyses – a pre-cultural, embodied given that exists independently of any discourse about it (HS1, 151). Foucault responds by suggesting that sex is “a complex idea formed inside the deployment of sexuality” rather than being something in reality that anchors the discourses of sexuality (HS1, 152). As the technology of sexuality developed, “sex” came to be posited as a thing in its own right “other than bodies, organs, somatic localisations, functions, anatomo-physiological systems, sensations and pleasures” (HS1, 152-3).

Whilst playing different roles in different domains, and referring to different entities, the concept of “sex” established an “artificial unity” by lumping heterogeneous elements together. This “fictitious
“unity” enabled “sex” to serve as “a causal principle, an omnipresent meaning, a secret to be discovered everywhere” (HS1, 154). Foucault is suggesting that sex, an apparently empirical concept, lacks clear empirical criteria yet operates as an explanatory principle for diverse phenomena. He is not denying the existence of bodies, organs, functions, pleasures, sensations etc, but rather that the concept “sex” picks out a determinate object without distorting it, manufacturing a pseudo-object like the hysteric’s “trauma”. Rather than being some real biological or psychological thing, sex is “the most speculative, most ideal, and most internal element in a deployment of sexuality organised by power” (HS1, 155).

Sex serves to link the knowledge of sexuality - in sexology, psychiatry, psychoanalysis and psychology - to the biological sciences and, without making any real use of these sciences, granting sexuality “a guarantee of quasi-scientificity” through this proximity (HS1, 155). Our self-understanding is now bound up with “what was [once] perceived as an obscure nameless urge” requiring us to expend enormous effort and employ great care in coming to know it” (HS1, 156). What is at stake is that if the concept sex provides an authoritative scientific foundation, as the ‘true’ foundation of our identities and desires, then it seems legitimate to control behaviours on the axis of normal-abnormal by their containment, control and channelling.

Foucault’s nominalism and anti-essentialism is in play here. He is particularly resistant to any move to naturalise human nature, to dictate some essential biological determinant of what we are. It is not however, that Foucault imagines that we are unconstrained by biological determinations. Rather he is cautious of the authority of science which hypostatises its objects as already given. This is particularly problematic for the human sciences in which man is the subject of knowledge of himself as an object. Given what I have discussed as the looping effects of power/knowledge in relation to interactive kinds, Foucault is right to be wary of any deterministic concept suggested by an epistemology that naively assumes that nature, including our own, simply imprints itself on our receptive souls. What Foucault’s genealogies bring into view is that our epistemic activity is a social activity shot through with power relations and that if these relations were to shift, knowledge would realign towards differently constituted objects. To see this is to realise that things could be different in ways we never expected. It is a means to prise open a space for starting to think.

**The Normalisation of Society**

Power not only produces subjects, objects and concepts but entire bodies of scientific know-how that feed back into society, restructuring it under the banner of technological progress. In SD, Foucault offers a genealogical account of the institutionalisation of technological knowledge which reflects this two-way process. Technological institutionalisation was a significant normative project, “one of the great instruments of power at the end of the classical age” (DP, 184). The emergence of technological
knowledge involved a battle “not between knowledge and ignorance”, but between “knowledges in the plural … defined by geographical regions, by the size of the workshops or factories and so on” (SD, 179). They struggled against each other in a society where knowing the secrets of technology was a source of wealth.

In this context of struggle, there developed processes that allowed “bigger, more general, or more industrialised knowledges, or knowledges that circulated more easily, to annex, confiscate, and take over smaller, more particular, more local, and more artisanal knowledges” (SD, 179). The state weighed in, selecting or eliminating certain knowledges, introducing norms enabling knowledges to be compared and to communicate. A hierarchical classification allowed knowledges to be fitted together from the most general to the most particular, whilst centralisation allowed surviving knowledges to be controlled, ensuring their selection, transmission or elimination. Accompanying this normalising, hierarchisation and centralisation, were a series of practices, projects and institutions - schools, curricula, standards, inspectors, journals, institutes, and so forth. Disciplinary power is applied to knowledges that become modern technological disciplines, each with an internal organisation, possessing criteria of selection (SD, 181). Knowledges formed disciplines with clear domains of inquiry and methods, organised into a hierarchy within an overarching discipline called “science”. Philosophy lost its foundational role and science, understood as a general domain, takes over the task of policing of knowledges.

From the 18th century, anything born or shaped outside the university had to prove itself on this institution’s terms. The amateur scholar ceases to exist. The tighter relationship between power and knowledge gives rise to a new constraint: “no longer the constraint of truth, but the constraint of science” (SD, 185). Foucault suggests that statements previously were considered on their merits, excluded or included on the basis of their content, “their conformity or nonconformity to a certain truth.” After the establishment of the modern disciplines, control is exercised over “the regularity of enunciations … who is speaking, are they qualified to speak, at what level is the statement situated, and to what extent does it conform to other forms and typologies of knowledge?” (SD, 184) Claims are judged in terms of the speaker’s or journal’s scientific authority, or the consistency with other scientific knowledge. Granting science this authority allows science to build knowledge without having to continually return to basic questions.

By portraying a battle between multiple knowledges, Foucault’s account enables us to see how technological norms, saturated with power/knowledge, were constructed. Canguilhem suggests that such technological norms should be seen within a broader normative project amounting to something
analogous to a system. Yet in any given historical society, the formation, operation and maintenance of norms as part of a systematic whole are beyond the conscious decisions of subjects. The formation of this social system of norms does not show up in conscious deliberations. Rather than the result of either choice or coercion, the technological “normalisation” of society involves the unconscious internalisation of what becomes a naturalistic background, developing by its own autonomous logic.

This scientific-technological normalisation should be seen as nested within Foucault’s analysis of modernity as the “carceral archipelago” (DP pp 293). Industrial technological normalisation isn’t distinct from the social normalisation Foucault calls discipline, or what we will shortly discuss as biopower. Techniques of industrial normalisation, such as the Fordist production line, presuppose reciprocal techniques of normalisation of workers who must internalise the norms required to render them docile and useful. The cars churned out on production lines then make suburbs possible, which then make cars necessary which remould subjectivities behind our backs.

Foucault’s argument that human bodily capacities are produced by disciplinary power challenges the conception of modernity which sees labour-power as simply given by nature and then taken up by capitalism’s various mechanisms. If labour-power or any human capacity is not simply given, we must account for its production. The “carceral archipelago” specifically refers to the proliferation of norms which mould subjects, as they internalise these norms. This process is backed by the authority of the human sciences, which are further refined by feedback from techniques of examination, observation, documentation, judgement and confession, to produce finer distinctions within knowledge and more precise applications of power. Foucault tell us that the supervision of this proliferation is firmly backed by medicine, psychiatry and their derivates, providing a veneer of “scientificity” (DP 296).

This carceral mechanism establishes a continuous gradation passing seamlessly from a transgression of the law to a departure from a rule or an average (DP, 298). “We are in the society of the teacher-judge, the doctor-judge, the educator-judge, the social worker-judge; it is on them that the universal reign of the normative is based; and each individual, wherever he may find himself, subjects to it his body, his gestures, his behaviour, his aptitudes, his achievements” (DP 304). Modern technologies of the self nurture a self that recognises itself in the normative categories of psychological and psychiatric sciences and their associated disciplines. Students, patients, factory and office workers are all made subjects of modern power for which there is no centre, no teleology, but many small unco-

---

67 For example, the standard 1.44 meter rail gauge established a norm reflecting the best compromise between the conflicting requirements of mechanics, fuel, trade, economics and politics, ultimately reflecting “an idea of society and its hierarchy of values.” (Canguilhem 1999, 247)
ordinated causes. Foucault’s account argues against global liberation. All politics must be local, even personal. We will see in the next Chapter that the struggle against normalisation involves a mode of fashioning an ethical way of being a self.

5 BIO-POWER and GOVERNMENTALITY

In the final lecture of SD, and later the same year, with the publication of HS1, Foucault broadens his analysis to take up the notion of “bio-power”. Biopower bears on populations. Discipline operates on individual bodies by constituting physical routines and structuring perceptual and social grids. At times Foucault describes discipline and biopower as two distinct (though overlapping and complementary) forms of power (e.g. in STP, lectures 1 and 2). Elsewhere he includes discipline within biopower, as one of the two levels at which biopower operates.\(^68\) However we characterise the distinction, biopower and disciplinary power are necessarily entwined, since bodies make up populations. Discipline is not a form of individualisation that is applied to already existing individuals, but rather it presupposes a multiplicity. Populations constitute disciplined individuals into a new political form that exceeds their mere combination.\(^69\) In fact, the disciplinary institutions are indispensible to the monitoring, and even more so to the interventions, of biopower. Biopower involves the administration of norms of the population as a whole, in order to regulate mass phenomena such as “the problems of birth-rate, longevity, public health, housing and migration” (HS1, 140). Whilst discipline emerged in institutions from which it diffused, biopower is primarily initiated through the state which seeks knowledge of the population’s characteristics, structures and trends so that it can be managed or compensated for what cannot be controlled.

The appearance of biopower, in the mid-19th century, must be seen in the context of modernity, with its medical and scientific knowledge and organisation of production in large urban centres, accompanied by poverty, housing shortage, poor sanitation, crime and disease. Biopower developed in tandem with sciences of demography, statistics, epidemiology, the analysis of the circulation of wealth and the evaluation of the relationship between resources and inhabitants (HS1, 140-1).

\(^68\) In HS1, Foucault argues that “this power over life evolved in two basic forms; these forms were not antithetical, however; they constituted rather two poles of development linked together by a whole intermediary cluster of relations.” (HS1 P. 139) The disciplinary pole is “centered on the body as a machine; its disciplining, the optimisation of its capabilities, the extortion of its forces, the parallel increase of its usefulness and its docility, its integration into systems of efficient and economic controls.” The biopower pole focussed “on the species body, the body imbued with the mechanics of life and serving as the basis of biological processes.” (HS 1 P. 139)

\(^69\) Biopower “does not exclude disciplinary technology, but it does dovetail into it, integrate it, modify it to some extent, and above all, use it by sort of infiltrating it, embedding itself in existing disciplinary techniques.” SD (242)
Biopower is increasingly incorporated into a continuum of apparatuses (medical, administrative etc) to regulate deviations from a norm.

From the 18th century, a “relative control over life” opened up “methods of power and knowledge [which] assumed responsibility for life processes and undertook to control and modify them. Western man was gradually learning what it meant to be a living species in a living world, to have a body, conditions of existence, probabilities of life, an individual and collective welfare, forces that could be modified, and a space in which they could be distributed in an optimal manner” (HS1, 142). This understanding of biopower was foreshadowed in terms of the discursive shifts charted in both OT and BC. In OT, Foucault outlines how the concept of life provided the basis for an understanding which depends not on surface phenomena, but on functional elements hidden from view. Living species become understood in terms of the particular manner in which each is linked to its environment (OT, 298). This understanding of individual bodies in terms of an environmental milieu is reflected in biopower’s ambitions to foster life at the level of the population, in which each individual is part of an environment for every other individual. The norms of modern medicine contribute to the human sciences and their deployment by biopower. Conceiving life as the “bipolarity of the normal and the pathological”, medical knowledge was directed by biopower towards the “benefit of the state” or “the order of a nation, the vigour of its armies, the fertility of its people, and the patient advance of its labours” (BC, 35). Biopower targets the population characterised by its own processes and laws and phenomena that include not only disease but production of wealth and its circulation. It regulates by “establishing a sort of homeostasis … achieving an overall equilibrium that protects the security of the whole from internal dangers” (SD, 249). The development of biopower is also bound up with how Foucault understands liberalism as a way of governing life such that it can be controlled by disciplinary forces (BB, lectures 11, 12).

Biopower doesn’t merely extend existing power, but represents an epochal transformation, “undoubtedly one of the most important in the history of human societies” (Foucault, cited by Senellart in STP, 369). By harnessing science, biopower reconfigures political action and determines its goals in terms of the life of the population. Foucault tells us that it was biopower that “brought life and its mechanisms into the realm of explicit calculations and made knowledge-power an agent of transformation in human life” - implying that knowledge-power did not previously have this role. Here Foucault is recasting his thesis on the birth of Man. In OT, he tells us that Man emerged at the dawn of modernity as a transcendental subject who knew himself as an empirical object. Biopower is embedded in this empirical knowledge. The human sciences take man as their object, understood as a living species within its environment. “If the question of man was raised – insofar as he was a specific living being, and specifically related to other living beings – the reason for this is to be sought in the new mode of relation between history and life: in this dual position of life that placed it at the same
In the modern era, knowledge produced by the human sciences is harnessed by the pastoral state. Biopower is clearly a two-edged sword. If humanity is a thing that can be made, Foucault shows us that it is not made with rational oversight or consciousness, but is the product of many small, forgotten acts of conditioning and resistance. By actions directed at populations, humans might succeed in mitigating the vicissitudes of precarious existence. But the nature of science’s authority worries Foucault. By this authority, not granted by “science itself” but by its connections to other discourses, scientific knowledge presupposes the goals and purposes of life, thus closing questions which he would want to remain open. He is concerned that biopower’s abstract, risk-averse imperative to optimise life, underpins its right to kill, which is “simply the reverse of the right of the social body to ensure, maintain, or develop its life” (HS1, 136). By exclusionary racism, war, eugenics or euthanasia, biopower ensures, by its “explicit calculations” that the life of the population, in an abstract sense, comes out ahead.

Foucault tells us that resistance to biopower is everywhere, though this resistance has not sought a return to sovereign power. Instead it relies for support on what biopower has invested in - man as a living being. Biopower’s notion of life has been “taken at face value and turned back against the system that was bent on controlling it”. By the tactical polyvalence of discourses, resistance takes the form of demands for the “right” to “life, one’s body, health, happiness, to the satisfaction of needs, and beyond all oppressions and ‘alienations’, the “right” to discover what one is and all that one can be” (HS1, 145). According to Foucault, these are not rights in any classical juridical sense, but rather struggles framed around the question of life. Just as the sexual liberationists are still caught within the device of sexuality, so those who resist biopower frame their resistance in the terms it dictates.

SD and HS1 raised the question of biopower in a cursory and generalised form. In the 1978 and 1979 courses, Foucault discusses biopower not in terms of the biological and physical life of the population, but within a broader framework of “governmentality”, a specific political rationality arising in the 18th century. These lectures can be seen a conceptual hinge by which Foucault turns away from the Nietzschean discourse of battle to the problematic of the “government of self and others”, the concern of his lectures in the 1980’s. In the 1980 course, the notion of “government” explicitly articulates “the
broad sense of techniques and procedures for directing human behaviour. Government of children, government of souls and consciences, government of a household, of a state, or of oneself” (EW1, 81). The question of how one should conduct oneself is clearly an ethical question, implicit in these lecture courses, becoming thematic in Foucault’s work of the 1980’s. In his genealogical work, Foucault insisted that knowledge and power could not be analysed separately. Now armed with this concept of governmentality, he will be able to bring a positive notion of subjectivity into this constellation.

Governmentality, guided by political economy, is directed towards the population, which appears as “a thick natural phenomenon”, opening up a new field for power and knowledge (STP, 71). Within this broader context, Foucault explores connections to liberalism and neo-liberalism. He sees liberalism not in terms of economic or political theory, but as the art of governing that submits political decision-making to the judgement of the market, constituted as a natural site of veridiction by the scientific discipline of political economy.70 Foucault’s genealogy charts how liberal governmentality emerged as a form of power/knowledge. By analysing society in terms of its specific nature, economists reveal it as a field of objects, amenable to knowledge and intervention. This natural order must be investigated since it dictates the basis of government actions which are seen to depend on internal regulation of a social mechanism that largely regulates itself. The market appears as a natural, spontaneous mechanism that one has to discern and follow, since attempts to modify its mechanisms will generally impair and distort them. Rather than impose injunctions or imperatives, the state’s role will be to manage these natural processes (BB, 31).

Political economy is not knowledge of government itself, but rather a science which is “tete-a-tete with the art of government … external to the art of government and that one may perfectly well found, establish, develop, and prove throughout, even though one is not governing or taking part in this art of government” (STP, 351). What follows is that political economy appears value-free, a science open to the critical scrutiny of all. But it is also indispensable to government with which it is really tete-a-tete, in other words, deeply entwined and complicit. Foucault’s further comments support this interpretation, suggesting that political economy is a particular amalgum of science and government, a “kind of more or less confused magma … an art of government that would be both knowledge and power, science and decision … two poles of a scientificity that on one hand, increasingly appeals to its theoretical purity and becomes economics, and, on the other, at the same time claims the right to be

70 By positing the market as a spontaneous natural mechanism, political economy is linked to biology’s notion of an inner organisation characterised by life as a dynamic and abstract principle common to all organisms (OT, 248). The ensuing biological notions of self-regulation and self-preservation are echoed in liberal conceptions of market freedom and autonomy assumed by political economy.
taken into consideration by a government that must model its decisions upon it” (STP, 351). This theme is seen throughout Foucault’s genealogy - the scientific discourse which appeals to its purity whilst still entwined with power, resulting in a “confused magma” which blurs the distinction between social facts and prescriptive norms. By such entwinement, political economy assumes an authoritative position as a science to endorse the rationality of markets as though such rationality were rationality *per se*. Whilst appearing to be just delivering knowledge, science results in dysfunctions such as the freedom-inhibiting normalising consequences that Foucault’s genealogies point to.

Foucault grants the objects of governmentality, such as population and the market, an ontological status somewhat similar to other aspects of social reality, such as sexuality, delinquency or madness. “The point of all these investigations concerning madness, disease, delinquency, sexuality and … [governmentality] is to show how the coupling of a set of practices and a regime of truth form an apparatus of knowledge-power that effectively marks out in reality that which does not exist and legitimately submits it to the division between true and false” (BB, 19). To make sense of this, we need to consider how such concepts are constructed within a context of power. The liberal concept of “market” is, for example, linked to concepts such as “nature” which lend it legitimacy. It has properties exceeding a mere aggregation of exchanges and, like the concept “sex” serves a range of functions. Once recognised, it is objective but not independent of this recognition. It possesses the looping effects of Hacking’s interactive kinds by which description influences phenomena which loops back to influence the description. Objects such as markets or delinquents “are not things that exist, or errors, or illusions, or ideologies. They are things that do not exist and yet which are inscribed in reality and fall under a regime of truth dividing the true from the false” (BB, 20). Foucault is saying that these objects don’t exist like the mind-independent objects of natural sciences. But once picked out by concepts and given a role in scientific theories, these objects are constituted as reality, as candidates for true statements which play a role in social practices.

6 NORMATIVE CONFUSIONS

I began this chapter by noting that science is conventionally understood as a form of knowledge that develops methods to abstract from the particularities of interests, prejudices and power in order to reach universal truths. In concluding, I will return to this theme, in the light of Foucault’s notion of power/knowledge and Habermas’s criticism that Foucault has an implied justificatory framework which relies on tacit appeals, presumably to values like Kant’s normative notions of dignity and autonomy. Foucault can’t own up to this framework because it is ruled out by his totalising critique of modernity. He commits a performative contradiction and stands accused of *crypto-normativism* - the smuggling in of values in contradiction to his value-free stance (PDM, 284). According to Habermas,
Foucault’s stance requires him to bracket any evaluative judgement of practices, leaving him no position from which to evaluate regimes of power/knowledge. Yet Foucault consistently assumes a critical position in relation to modern society, a position of “arbitrary partisanship of a criticism that cannot account for its normative foundations” (PDM, 276). On Habermas’ reading, Foucault argues that all claims are relative to their contexts which he describes by reducing the internal aspects of meaning, validity and value to power, so denying himself any normative position to support his condemnations. If genealogy is no more than historical descriptions, gratuitously embellished with unfounded judgements, his “critique” of the human sciences loses any bite. Nancy Fraser agrees asking “why is struggle preferable to submission? Why ought domination be resisted? Only with the introduction of normative notions could he begin to tell us what is wrong with the modern power/knowledge regime and why we ought to oppose it.” (Fraser 1989, 29)

Michael Kelly defends Foucault by suggesting that normative premise is there all along. Freedom is presupposed and not undermined by power (Kelly 1992, 365-400). Another response is to refuse the terms of engagement of Habermas’ argument by suggesting that Foucault’s critique is directed towards the very idea of normativity (Tulley 1999). Rather than tease out the insights and shortcomings of these defences, I will argue that Habermas’ criticisms misunderstand Foucault’s project. I will develop this in Chapter 8 where I argue that Foucault’s work is prescriptive and has a normative basis, although not of a sort discernible by conventional normative grammars.

It is in DP, that we see Foucault’s most explicitly evaluative language in his condemnatory accounts of the “carceral archipelago” with it “insidious leniences, unavowable petty cruelties, small acts of cunning …” (DP 308). But it is not simply injustice or cruelty that Foucault evokes in DP. He wants to reveal the present situation in which we are trapped because we can’t see it clearly. By suspending normative and epistemic commitments, he can reveal the workings of power/knowledge. DP opens with an account of a horrific execution, taken from contemporary records. Foucault immediately proceeds to present a prison timetable, again from contemporary documents. Having abstained from any commentary or judgement, he then drily remarks, “We have then, a public execution and a timetable” (DP, 7). What Foucault wants us to see here is not the obvious cruelty or injustice of the punishments (it is assumed that we do see this) but the play of power/knowledge, the constructions of categories of thought that allow reasonable people to accept these regimes, and consequently, how his readers’ deepest convictions of what is natural and obvious are similarly constructed. Likewise, in his genealogies of sexuality, Foucault doesn’t emphasise the suffering, or champion the rights, of marginalised sexualities, but is concerned primarily with the larger theme of the fabrication of what we take in human nature to be most natural. Foucault’s genealogy doesn’t prescribe how we should think, but aims to open up the space in which we can think.
In a late essay, Foucault urges us to refuse the “blackmail” of the enlightenment – the demand that one must be for against enlightenment (PT, 109). He wants to undertake “rational criticism of rationality”, carried out from a position which doesn’t pronounce on the “correct” rationality (as though from outside all conditioning) but only points to rationality’s ambiguity, indispensability and, at the same time, its dangers. He tells us that Social Darwinism was “an irrationality … that was at the same time, a certain form of rationality.” Philosophy’s function is to “accept this sort of spiral, this sort of revolving door of rationality that refers us to its necessity, it indispensability, and, at the same time its dangers” (EW3, 358). Foucault’s diagnosis suggests modernity denies rationality’s ambiguity.

In the last Chapter, I referred to Foucault’s diagnosis in terms of Koopman’s notion of eliminating ambiguity by purification (Koopman 2002). Not only madness and reason, but power and freedom, or power and reason, are pairs that are “both reciprocal and incompatible” (HM, 529). Reason can neither master madness, nor can it detach itself from it. Scientific modernity institutes a division to preserve each in its purified form, amounting to the simultaneous production of these pairs in ways that neither can admit admixture with the other (Koopman, 157). Against such scientific modernity, Foucault has always been concerned with establishing dialogue across categorical borders, mixing the concepts that modernity insists must remain pure - reason and madness, power and freedom, power and truth. Han interprets this situation as an extreme ambivalence on Foucault’s part, such that on the question of freedom he is “torn between two irreconcilable extremes” - the subject as an autonomous free actor and the subject defined by a set of power relations (Han 2002, 172). But if this fundamental ambivalence is not Foucault’s, but the problematic of modernity itself, then this must change how we see what Fraser calls his “normative confusions” (Fraser 1981). Foucault is not primarily showing us the truth, falsity, rightness or wrongness of madness, freedom, reason, power, or delinquency. Rather he is trying to open dialogue across boundaries of concepts which define themselves in opposition to each other, to show us that madness is not always madness, freedom not always freedom, that reason has a history, power doesn’t negate freedom and delinquency was invented in the 19th century.

Despite his bleak account of the carceral archipelago, Foucault’s work need not be read as total condemnation, leaving him nowhere to stand. He wants to maintain the tension of ambiguity by showing that we are both conditioned by power and free and reasonable.

Habermas’ criticisms reflect his misunderstanding of Foucault’s intentions as attempting to provide genealogy with “a scientific status [that] will someday be comparable to that of the natural sciences” (PDM, 275). Foucault explicitly locates his own reflection as an attempt to break free of the epistemic sovereignty ascribed to science, which involves the adoption of a position, as if outside its historical and social situation, from which truth claims can be validated. Instead, Foucault acknowledges his social and historical embeddedness, offering a conception of power and knowledge without any sovereign standpoint. He is not pulling the rug from under his feet by claiming that all claims are
tainted with power and therefore illegitimate. Rather he is claiming that legitimating reasons, his included, will always be caught up with the power relations of their particular situations and to lose sight of this is dangerous.

Given this position, it is not clear if, or why, it is a problem that Foucault describes modern society in evaluative terms whilst withdrawing from any explicit normative foundational framework or criteria. Rather than simply judging certain practices, he is problematising their authoritative scientific underpinnings, drawing them out from an unquestionable background to open the possibility of thinking differently. Genealogy discloses contingent aspects of practices which then open themselves to evaluation, though not in terms of a prior framework, located outside the situation, but as historically and socially embedded responses to particular situations. The evaluative aspect of Foucault’s work does not flow, pretend to flow, or need to flow logically from his descriptions of contingent constellations of power/knowledge or, for that matter “normative foundations”. Evaluation is undertaken by an imaginative response which enlists resistances to power, drawn from Foucault’s and the reader’s unarticulated backgrounds of experiences, which can’t be captured by a general framework.

Foucault’s genealogies allow us to recognise how objects - the delinquent, the hysterical, the homosexual - are constructed by particular alignments of power/knowledge, not simply discovered as something that was always there. As the most authoritative form of knowledge, it is science, particularly human sciences, from which these power effects flow. Foucault wants to make us aware of this and how they could be otherwise. His critique is “not a matter of saying that things are not right as they are. It is a matter of pointing out on what kinds of assumptions, what kinds of familiar, unchallenged, unconsidered modes of thought the practices we accept rest” (EW3, 456).

I will return to the question of Foucault’s normative foundation in Chapter 8, where I will articulate the fundamental intuitions which motivate both Foucault’s and Habermas’ projects by placing them in a broader framework. In this chapter, we have seen the inseparability of power and scientific knowledge, how particular sciences have arisen within particular configurations of power which, rather than merely distort, produce and link bodies, concepts, discourses, practices, objects and subjects, to form chains that relay power/knowledge. In the following chapter I will turn to Foucault’s incorporation of the subject into the power/knowledge constellation by ethical practices and consider how Foucault regards philosophy as not like science, a problem-solving activity issuing in determinate though fallible answers, but a way of life, an ongoing activity of questioning in itself.
Chapter 4
Ethics

Thus far we have seen Foucault concerned with limits prescribed either by the archaeological *episteme* or genealogical power/knowledge. By bringing these limits into view, he wants to open up the possibility of a perspective beyond them. His critics have pointed to the difficulties in locating himself in relation to these limits (he seems trapped within the limits he identifies) and articulating his own normative standpoint (he provides no reasons apart from his own preferences for surpassing these limits). In the following chapters, we will see Habermas harness science to help articulate him a normative framework. But Foucault, acutely aware of science’s inextricable entwinement with power, needs a broader framework, into which science sits as one form of rationality amongst others. His genealogies of the subject respond by challenging the authority modernity grants science as ultimate arbiter of knowledge. By revealing the contingent emergence of scientific thinking, eclipsing earlier modes of relation to the self, science comes into view as one regime of truth, alongside other possible regimes.

Importantly, Foucault is not judging science to be “true or false, founded or unfounded, real or illusory, scientific or ideological, legitimate or abusive” but is revealing “the connections that can be identified between mechanisms of coercion and elements of knowledge … such that a given element of knowledge takes on the effects of power in a given system where it is allocated to a true, probable, uncertain or false element, such that a procedure of coercion acquires the very form and justification of a rational, calculated, technically efficient element etc.” (PT, 59). He will call this endeavour “critique” so aligning himself with the critical theory tradition. The primary task of Foucauldian critique is not to evaluate knowledge or power, but to bring into relief the framework of evaluation itself. Critique inquires into the particular relations of knowledge to power that result in our epistemological certainties supporting a way of structuring the world that forecloses alternative possibilities.

Foucault’s articulation of this broader perspective is enabled by his shift to a more active and responsible subject, yet still accommodating his previous perspectives. Whilst this shift was gradual, the 1980 lectures suggest a decisive commitment towards Foucault’s final period of his work. The four following lecture series from 1981 to 1984, and the two further volumes of the *History of Sexuality*, consolidate this as a perspective from which he doesn’t deviate. Starting from the Socratic dialogues, Foucault’s genealogy of the subject traverses ancient philosophy, the Christian pastoral to the “Cartesian moment”, the decisive break that opened up the modern scientific worldview. This
rupture is consolidated by Kant’s transcendental philosophy and the rise of Man as the modern subject-object of science.

As we will see, this genealogy allows a problematisation of science, as a form of knowledge which is relative to a particular form of subjectivity. Foucault’s three constitutive axes of the subject – knowledge, power, and self-constitution – are now inextricably entwined. But rather than abandoning power and knowledge to talk about ethics, Foucault complicates these aspects by including the subject which becomes political as both locus of subjugation to, and resistance against, power/knowledge, particularly scientific power/knowledge. In this chapter, I will firstly consider Foucault’s turn from power to the subject in the 1980 lectures and whether his new focus on agency, autonomy and freedom is consistent with his earlier work. I will then analyse Foucault’s notion of “regimes of truth”, a concept that both reveals how Foucault views science, and sets up his genealogy of the modern subject in terms of its constitution by truth, power and knowledge. I will consider Foucault’s concepts of ethics, aesthetics and spirituality as a basis for discussing this genealogy. By tracing the subject’s constitution from antiquity to the post-Enlightenment, we will locate Foucault’s view of science within modernity in relation to critique and philosophy.

1 FROM POWER TO THE SUBJECT

The title of the 1980 lecture series On the Government of the Living (hereafter GOL), suggests an intention to return to biopolitics, bringing to bear reflection enriched by Foucault’s work on governmentality. However, Foucault had been visiting the United States and was aware of the criticism of his treatment of the subject as a passive object of power. His 1980 course therefore responded by leaving behind earlier themes of biopower and the government of living populations. The “government of men by truth”, gives effect to the shift from the “worn and hackneyed theme of knowledge-power” (GOL, 11). This concept of government now includes the subject - governing, governed and self-governed. The notion of “government by truth” involves the subject in its relations to the manifestation of truth, or the ways in which subjects bind themselves to truth.

As long as Foucault was studying the modern West, he treated the subject as the product of systems of power/knowledge. His archaeologies displaced the subject by discourse, his genealogies displaced it by power. Prompted by his reading of ancient philosophy Foucault became increasingly aware that the

---

71 See discussion of biopower and governmentality in Chapter 3.
72 Szakolczai (1998) notes the importance of discussions with Dreyfus and Rabinow on Foucault’s 1979 visit.
73 Foucault wants to “[to get] rid of knowledge/power” which he tells us he had adopted to avoid the shortcomings of the notion of “dominant ideology” which assumes that power and knowledge are antinomic. (GOL p.12)
subject could not simply be analysed as an object subjected to power/knowledge, but must also play an active role in its own shaping. Foucault’s genealogical works had bracketed the subject, to portray power as not located, but comprising a vast network of anonymous relations, relayed through individuals. Now, with “government”, ethics comes into focus since one must govern oneself and others. With ethics, the constituted subject, still constituted by power/knowledge, turns back upon itself to critically examine the processes by which it was constituted.

Foucault’s interest in ethics also arose from his awareness of the current political situation, which he analyses in terms of governmentality. He identifies new social struggles as responses to the modern state taking over the pastoral role of the church and governing citizens by scientific norms. These struggles are “not exactly for or against the ‘individual’, but rather … struggles against the ‘government of individualisation’” (EW3, 330). Against the “secrecy, deformation, and mystifying representations imposed on people” these struggles resist identities thrust upon individuals by use of administrative and scientific categories (EW3, 330-1). Foucault is referring to various contemporary movements such as gay liberation, black liberation, indigenous land rights, feminism and so forth and their resistance to scientific power/knowledge which “categorises the individual, marks him by his own individuality, attaches him to his own identity, imposes a law of truth on him which he must recognise and which others have to recognise in him” (EW3, 331). The struggles are not about the truth or falsity of scientific knowledge, but “the way it circulates and functions, its relations to power. In short, the régime du savoir.” Resistance is directed against subjection (explicitly in the sense of making individuals into subjects) by administrative-scientific knowledge-power.

Some critics think Foucault’s genealogical accounts of pervasive power are incompatible with this focus on self-reflexive practices, which they think implies autonomy and independence from power. McCarthy claims that genealogy treats “the subject merely as an effect of power” and then goes on to assume that Foucault thinks that the subject is merely the effect of power. He concludes that this means that if Foucault is right about power, he has no basis for the subject’s autonomy. Alternatively, if subjects are autonomous, then his account of power is wrong (McCarthy 1992, 258). As far as I am aware, Foucault has been careful to never say that the subject is merely an effect of power or discourse. He has however, frequently said that the subject is an effect of power or discourse, intending to convey the extent to which subjects areconditioned by contingent arrangements of power/knowledge. What we see in his ethics is an option that has always been available to him, to correct what he now recognises as an over-emphasis by highlighting the moment of autonomy within action, even though that autonomy is never without constraints (EW1 281-2).

Foucault never intended to deny the role of the subject but to account for its constitution within a historical framework (PK, 117). In his 1980 lectures at Dartmouth College, he discusses his current
work as “a genealogy of this subject, [which studies] the constitution of the subject across history which has led us up to the modern concept of the self” (Foucault & Blasius 1993, 202). By means of this genealogy, he can avoid treating “the foundation of all knowledge and the principle of all signification as stemming from the meaningful subject” (Foucault & Blasius 1993, 201). Rather than foundational and fixed, Foucault’s subject is relational and dynamic, actively becoming (EW1, 290-1). Foucault discussed two “technologies”- a technology of domination and a technology of the self (Foucault & Blasius 1993, 203). Later that year in Berkeley, he admitted “when I was studying asylums, prisons, and so on, I insisted too much on the techniques of domination” (cited in McCarthy 1992, 279 n.64). Now his task is to correct this over-emphasis by bringing technologies of the self into play, and treating power as “due to the subtle integration of coercion technologies and self technologies.” This “subtle integration” suggests that the two technologies are entwined in a way that both are present as moments in action. Just as the emphasis on power and discourse in Foucault’s earlier work bracketed but didn’t preclude the subject, now the self-reflexive subject does not preclude power.

Autonomy is perhaps best understood as not an absence of power but a movement of power away from normalisation. On this understanding, autonomy would entail a change in the relationship between these two technologies such that the subject is able to step back from normalising power. This process of “desubjectification”, requires “new forms of subjectivity through the refusal of this kind of individuality which has been imposed on us for several centuries” (EW3, 336). Neither the subject nor power is done away with. Rather a new form of subjectivity is constituted that is to some extent distanced from the normalising power both constituted and supported by the sciences. Foucault sees this process of distanciation paradigmatically in thought which he links to problematisation (EW1, 117). Thought problematises, disclosing the world in a new light, as what is given as an obstacle is transformed into a question (FL, 421). Foucault offers two ways to respond to power, two different ways of problematisation, corresponding to different levels of power. The first – ethics - aims at the micro-level of power located in a subject, whilst the second – critique - is directed at the level of the anonymous pervasive network of power relations he analysed in his earlier work.

Foucault’s shift in focus to the ethical notion of “care for oneself”, whereby the self acts on itself by shaping itself as a moral subject, no doubt prompted him to think of his own philosophical activity in

---

74 By “moments in action” I mean that one can “passively” accept power or can “actively” resist, but in either case, acceptance or resistance is an action. Yet, at the same time, the range of possibilities for such actions is circumscribed by a context of power that is not one’s own.
75 For definition of “problematisation” see Chapter 3, fn 6
76 I am indebted to Mark Kelly for the distinction between these two different analytical perspectives. (Kelly 2009, 61)
these terms. He characterises his work as “another kind of critical philosophy … that seeks the conditions and the indefinite possibility of transforming the subject, of transforming ourselves” (Foucault & Blasius 1993, 224 note 4). Such transformations apply to Foucault himself with his shifting perspectives. This was no post hoc justification but an insight prompted by Foucault’s understanding of his own activity in the light of what philosophy had been and still could be.77

As we have seen, Foucault resisted confinement within any one framework, seeking to remain receptive to new kinds of phenomena needing different frameworks (HS2, 6). He sought multiple standpoints to gain a broader view, each having its limitations which could only be overcome by a further shift. He doesn’t however, present the shift to the subject as just a new framework, but rather the organising principal of his entire previous work. He now claims that for twenty years, his objective has been “to create a history of the different modes by which, in our culture, human beings are made subjects” (EW3, 326). This constitution of subjects occurs within three historical dimensions represented by frameworks of discursive practices, power practices and practices of self on self. From 1980 onwards, in interviews and lectures, Foucault presents his overall project as revealing the contingent historical construction of the subject within these frameworks.78 None of these dimensions can be grasped independently since each mediates the others. This understanding of the subject opposes any foundational conception of subjectivity. Rather than possessing a universal and timeless essence discoverable by science, Foucault’s subject emerges in history, as it becomes the object of discursive practices, power and practices of the self (EW1, 318).

Foucault places his project within the history of philosophy, as part of a subterranean path of critical thought that he identifies as a way of life and traces through aspects of Plato’s *Laches*, the Cynics, the Stoics, De Sade, Kant, Hegel, Marx, Nietzsche and the Frankfurt School. He links the shifts in his theoretical framework to the self-transformation essential to the philosophical life, thus enabling him to situate his own reflection, something he previously had difficulty doing. Now with the subject not simply subjugated by power/knowledge but actively constituting itself, Foucault can see his own historic-critical reflection as a practice of the self, an attempt “to get free of oneself” (HS2, 8). This involves testing limits to explore what can be changed in oneself and one’s present. Whereas Foucault’s archaeologies and genealogies emphasise coercive subjection to scientific power/knowledge, ethics articulates a form of resistance. This resistance is not a rejection of science, or a denial of scientific truth, but a revaluation of science’s authority from which its power effects

---

77 McGushin plausibly argues that, after coming across the practices of the self, Foucault better understood his work as not merely diagnostic, but as *etho-poetic*, a term denoting the making of a character, an *ethos*. (HS, 237; McGushin 2007, xii-xiii)

78 see EW3, 326-7; HS2, 6; FR, 336-8; EW1, 262-3, 318.
flow. Foucault’s relativisation of science as one form of rationality amongst others is evoked by his notion of “regimes of truth”.

2 REGIMES OF TRUTH

In the 1980 lectures, Foucault talks of “regimes of truth” constraining individuals to certain “truth acts”, or roles taken on by subjects in the manifestation of truth. He asks in what ways, by what procedures, in view of what ends is a subject bound to the manifestation of truth? It is this relationship between the manifestation of the truth and the forms of the subject’s involvement that defines “regime”. Returning to the analytic of finitude (see Ch. 2), but now through the lens of ancient philosophy and its transformations through confession and uptake by the human sciences, Foucault poses the question “how has Western man bound himself to the obligation to manifest in truth what he himself is?” The genealogy that unfolds in these lectures examines “how this double bind, this regime of truth, by which men find themselves bound to manifest themselves as object of truth, is linked to political, juridical etc regimes” (GOL, 101).

The political connotations of “regime” are intended, and here Foucault anticipates an objection. It doesn’t seem right to talk about a “regime” of truth in the same way one speaks of a political, medical or penal regime (GOL, 94). Truth acts that require an obligation, such as believing in the resurrection, are not genuine truth acts, since they involve coercion of the non-true or the unverifiable. But in the case of genuine truth, there is no need of a “regime” of obligation. The coercive force resides within truth itself (GOL, 95). There is only need to speak of a regime of truth when something other than truth is involved. Foucault sees this objection as motivated by a view that truth is already there, waiting to be discovered when it will speaks to us directly without any need for coercion (PP, 235). Against this “philosophical-scientific standpoint”, Foucault recalls the archaic standpoint towards truth which it has covered up.79 The same truth does not appear for all, but depends on how the subject has been constituted. Foucault offers the example of Descartes’ “I think therefore I am” which seems “theoretically unanswerable” (GOL, 98). However, behind the explicit “therefore” Foucault locates an implicit “therefore” which issues from the acceptance of a particular framework, a regime

---

79 The archaic notion sees a “dispersed, discontinuous, interrupted truth” which only speaks in certain places to certain people. (PP, 236) It is a truth that is not waiting for just anyone to discover it, but a truth with favourable moments and propitious places, for example, the oracle at Delphi or the god who cures at Epidaurus. This notion of a truth-event contrasts to the universal ever-present truth of modern science, which presupposes that “there is truth everywhere, in every place and all the time.” (PP, 234) Such scientific truth may be difficult to reach but it is still there, waiting to be discovered. Nothing is ruled out in principle or necessarily hidden. (PP, 236)
of truth which requires a particular kind of subject who is “qualified in a certain way” to be compelled by evidence. Behind every scientific argument, reason or piece of evidence, there is a hidden prior commitment outside the framework of the regime. Whilst the rules of the regime dictate the division between true and false within the regime, these rules themselves don’t follow from a rule but require a commitment of a subject constituted in a particular way. As will be seen, Foucault reads Descartes’ Meditations as a way of constituting a particular regime of subjectivity that opens up the scientific worldview.

Despite the French language convention of referring to sciences in the plural, Foucault thinks that in terms of regimes of truth, it is legitimate to refer to science in the singular. “Science would be a family of games of truth all of which submit to the same regime … and this very specific, very particular regime of truth is a regime in which the power of the truth is organised in a way such that constraint is assured by truth itself” (GOL, 99). In other words, with science, the political nature of the regime recedes, defining itself as constrained by truth alone. The objectivating stance of science involves methods that cancel the contingency of the particular form of subjectivity required. This feature accounts for what appears as the imperialistic tendencies of science to reduce all values to scientific truth and dismiss what doesn’t fit within this schema as illusory. But from an “external” perspective, the regime of science is seen as one amongst other possible regimes. To overlook this perspective is to adopt the positivist assumption of an existing, though unknown, state of affairs towards which science progresses, and the non-truth of statements and non-reality of objects, outside science’s regime.

Foucault has long opposed such positivism. The ordering of the world into conceptual categories is an activity that precedes truth, since truth relies on such ordering for its subsequent establishment (WK, 208). But by assuming the existence of truth from the outset, a truth that knowledge gradually uncovers, positivism endorses science as the inexorable force to which one should stoically submit. Foucault wants to challenge this assumption and open the possibility of different divisions between the true and the false. In the 1980 lectures, he frames our activity of conceptual formation as a question of the constitution of subjects who are not merely subjugated objects, but also constitute themselves. This is ultimately an ethical question of how to care for oneself.

By his analysis of regimes of truth, Foucault eschews the binary opposition of science “in which the triumphant autonomy of truth and its intrinsic powers would reign” and “ideologies in which the false, or the non-true, would have to arm itself or be armed by a supplementary and external power in order to take on, improperly, the force, value, and effect of truth.” Rather there is a multiplicity of truth regimes, all of which, whether scientific or not, entail specific ways of linking the manifestation of truth to the subject. Whilst some regimes have a history and domain close to science, others are “quite
coherent and complex and very distant from scientific regimes of self-indexation of truth”. Science is not defined in opposition to ideology, but is “one among many possible and existing regimes of truth” (GOL, 100). Other regimes, for example religious, psychoanalytic or moral truths, require the constitution of different subjectivities in order to manifest (GOL, 99).

The notion of a regime of truth carries critical force, since it reveals the contingency of the scientific truth regime. There is nothing that necessitates, obliges or forces us to affirm the constitution of a subjectivity that demands a scientific stance that excludes all other stances. Foucault is not questioning the truth of scientific content. Rather he is challenging the authoritative value granted to scientific truth and the commitment to the belief that science imprints the world-in-itself directly onto our souls and doesn’t to some extent depend on us. Such commitments block the freedom of thought by freezing subjectivity, character or ethos, amounting to an ethical entrapment that Foucault wants to unmask.

3 ETHICS, AESTHETICS AND SPIRITUALITY

To draw out Foucault’s view of science and its relations to other practices such as philosophy, we need to examine the role his ethics in self-formation and resistance to scientific power/knowledge. Foucault sees the self as constituted by power relations, yet still capable of autonomous critical reflection and self-transformation. The subject exercises freedom by problematising her behaviour, beliefs and social context. This freedom is always mitigated. Whilst there is no overall liberation from power, local emancipations from particular systems of domination are still possible. Foucault’s ethics inquires into resistance, the possibility of contesting determinations by “refus[ing] what we are” (EW3, 336). Resistance to biopower involves “a refusal of these abstractions of economic, of economic and ideological state violence which ignores who we are individually, and also a refusal of a scientific or administrative inquisition which determines who one is” (EW3, 331). On one hand we resist by “refus[ing] what we are”, on the other hand, we must invent “new forms of individuality” (EW3, 336).

Foucault calls for a new type of ethics, not based on abstract philosophical ideas that obscure the complexity of specific concrete situations, but on the ethics of the self inspired by Greco-roman antiquity. This ethics is essentially a mode of self-formation not involving universal ethical codes, like Christian or Kantian ethics, that dominate modern Western society, permitting or prohibiting actions. In contrast, antiquity’s more individualistic ethics give an important place to self-

---

80 Foucault’s notion of ethics borrows the idea of practices of the self from the ancient Greeks, but not the actual contents of their ethics (EW1, 256).
formation. Care of the self aims at establishing the right relation with oneself, defined in terms of self-
mastery, tranquillity, harmony or joy (HS2, 28).

The care of the self implies care for others, since the relation to self forms part of a network of social
relations (EW1, 287). The care of the self isn’t a form of “inward-looking” attention. It is “a number
of actions exercised on the self by the self, actions by which one takes responsibility for oneself and
by which one changes, purifies, transforms, and transfigures oneself” (HS, 11). Such activities include
meditation, examination of conscience, checking representations which appear in the mind and so
forth. These activities are not ways of coming to know an object in scientific terms, but rather involve
monitoring, protecting, admonishing, curing, rewarding or cultivating oneself. By means of such
exercises, “the subject constitutes itself in an active fashion” (EW1, 291). This active self-
constitution, or “subjectivation”, contrasts to the self’s subjection to the disciplines or biopower.81 It is
the self’s own action on itself, striving for principled ethical coherence by undertaking ascesis, in
order to create an ethos, “a mode of being for the subject, along with a certain way of acting, a way
visible to others”82 (EW1, 286). Whilst Foucault refers to these practices as “practices of freedom”
chosen by an individual, they are “not something invented by the individual” but “are models that he
finds in his culture and are proposed, suggested, imposed upon him by his culture, his society, and his
social group” (EW1, 291). Whilst practices might be freely chosen, they are chosen from a pool of
possible choices given by society.

Foucault thinks the culture of care of the self still significant because it provides the resources to
challenge the normalisation of the modern subject by which we are compelled to seek the truth of
ourselves in science (HS 9; EW1, 255-6). Care of the self in antiquity was bound up with living a
beautiful, noble, and memorable life and was the framework within which one understood the
injunction to know oneself (HS, 4). But self-knowledge always involved prudence in a particular
setting. “Know thyself” (gnothi seauton) was subordinated to care for self (epimeleia heautou). To
know oneself meant to know one’s place, one’s limitations and was part of caring for oneself. It was
by the self’s primary relation of care to itself, that the self became constituted as an object of
pragmatic, not theoretical, knowledge. In modernity, the situation is reversed. The care of the self can
only take place through knowledge of scientific truths about the self. However, if the modern self
were to understand itself as something to be cultivated, it could subscribe to an individualistic and

81 “Subjectivation” is a term coined by Foucault in the early 1980’s to capture the idea of the active
constitution of oneself as a subject. See McGushin 2007, 304 note 6; Kelly 2013.
82 Ascesis is the practices through which one becomes a subject, producing a certain type of relationship that
one has with oneself. Ascesis does not necessarily require deprivation or renunciation.
voluntaristic ethics not requiring the support of the normalising power of tradition or science. As its own artist, the self would enjoy the autonomy that modernity requires.

Foucault emphasised that ancient ethics was not prescriptivist or universalist, but was directed at aesthetic qualities such as beauty or nobility (EW1 261). Like any skill, it was not a matter of learning facts, but a *techne* of artfully constructing, cultivating through repetitive exercises, a certain sensitivity within oneself to oneself. 83 However, Pierre Hadot thought that Foucault’s notion of a culture of the self was *too* aesthetic (Hadot 1995, 211). 84 He emphasises that the ancient care of the self is directed towards “the best portion of oneself” which is “in the last analysis, a transcendental self.” Ancient ethics is not concerned with the pleasure one finds within oneself as Foucault suggests. The happiness of an ethical life consists “in virtue itself, which is its own reward” (Hadot 1995, 207). Hadot thinks that the practices of the Stoics or Platonists not only have to do with a relationship to oneself but “the feeling of belonging … both to the whole constituted by the human community, and to that constituted by the cosmic whole” (Hadot 1995, 208).

Hadot’s concern seems to be that Foucault’s talk of ethics as aesthetics, stylisation and pleasure appears to suggest the reduction of ethics to pleasure. In response, I will firstly clarify one thing that Foucault was *not* talking about. He rejects “the Californian cult of the self” in which “one is supposed to discover one’s true self, to separate it from that which might obscure or alienate it, to decipher its truth thanks to psychological or psychoanalytic science, which is supposed to be able to tell you what your true self is” (EW1, 271). Foucault was very critical of the self-absorption and introspection of Western culture. It seems plausible that Foucault employed expressions such as “aesthetics of the self”, precisely to distance his notion from the “Californian cult” which assumed a reality knowable by science. An aesthetics of the self suggests the crafting of the self by the self, not according to something discoverable by science.

Foucault need not be held accountable for an accurate interpretation of ancient ethics, since this was not his intention. Rather he was inspired by, and borrowed key notions from, ancient ethics. But he also included the notion of aesthetics. In a Kantian sense, aesthetics is the experience of a common sense, defined by communicability and inseparable from the desire to communicate. Whilst aesthetic judgments are subjective, aesthetic reasoning aspires towards universality. But this doesn’t involve subsumptions under immutable universal categories. Yet it also isn’t relativism. There is always something at stake that, whilst not achieving universal consensus, still matters (Kant 1987, 159-62).

83 By *techne*, I refer to knowing how to do or make something, something that, like art, can’t be specified fully in advance due to its variability and exploratory nature.

84 Foucault’s embrace of ancient ethics was very influenced by reading the work of Hadot, a colleague at the *College de France*. see Hadot 1995.
We can also see Foucault’s notion of aesthetics suggesting resistance to scientific ontologizing. Foucault assigns a transgressive role to art, which can cut across categories of thought.\(^8^5\) Foucault’s ethics is aesthetic not by being oriented towards beauty but by regarding our lives as material to be transformed. In describing an ethics which is individualistic, voluntaristic, non-normalising and transgressive, it seems apt to include aesthetics, despite the fact that this inclusion fails to square with ancient ethics. Foucault’s interpretation of ancient ethics is directed to a present in which, as Hadot himself admits, “‘universal reason’ and ‘universal nature’ do not have much meaning anymore” (Hadot 1995, 208).

Ancient ethics and philosophy were structured by the relationship between subjectivity and the truth, which Foucault called “spirituality.”\(^8^6\) Whilst philosophy is concerned with knowing oneself, spirituality is “the search, practice and experience through which the subject carries out the necessary transformations on himself in order to have access to the truth” (HS, 15). The subject’s access to truth is grounded on ascesis, on transforming her mode of being as a subject by certain exercises. Spirituality is the price paid for truth which “is never given to a subject by right” (HS, 15). In order to gain truth the subject must undergo a sort of conversion, by transforming their way of being a subject, their way of seeing and inhabiting the world. Truth is not experienced as correspondence between beliefs and states of affairs, or propositions and experiences but comes about as salvation, happiness, tranquillity or fullness (HS, 16, 17). This stands in contrast to modern philosophy, where spirituality has been occluded and access to truth is by knowledge gained by method and evidence akin to science.

Foucault’s notion of truth is both diagnostic, since it discloses power/knowledge relationships, and etho-poetic, since it transforms the mode of subjectivity. This truth is not acquired by mere acts of cognition. The subject, as that which stands opposed to objects in the world, does not by its nature have access to the truth of those objects, or of itself. To modern ears, the notion of truth accessed through spiritual practices may sound like a relic of a superstitious past that we should put behind us. But Foucault maintains that the reason why spirituality is so difficult for us to grasp is due to a historical event in which the scientific knowledge of the self came to overshadow the care of the self. In modernity, we gain access to truth by scientific knowledge, as “access to a domain of objects” by objective, methodical thought, logical analysis, evidence etc. (HS, 191). The thought of spiritual access to truth, experienced as fulfilment and salvation, no longer has any meaning beyond religion or self-help seminars.

\(^8^5\) Boothroyd (1996) points out this transgressive dimension in art.

\(^8^6\) In Foucault’s use, spirituality doesn’t involve any commitments to religious concepts such as supernatural beings or immortality.
One can of course think of aesthetic, moral or religious experiences and agree that one must be constituted - formed, conditioned, indoctrinated, educated – as a subject in certain ways in order to have access to these experiences. But these are usually not considered experiences of “truth” in any literal sense. One way to grasp what Foucault is getting at is to consider psychoanalysis and Marxism, which he tells us are “not exactly sciences … but possess “at least certain elements, certain requirements of spirituality.” Two questions - what the subject must be in order to gain access to truth, and what aspects of the subject will be transformed by access to truth - are “absolutely typical of spirituality [and] are found at … the source and outcome of both these knowledges” (HS, 29). Foucault is suggesting that we have to be a certain way in order to grasp the truths of psychoanalysis or Marxism and that these truths will change us. Psychoanalysis and Marxism require a certain receptivity that is more than just attentiveness to facts or logic. The truth involved is not the correspondence of states of affairs to propositions. It involves a sensitivity that will enable us to experience what was totally unexpected and could not be deduced merely from an acquaintance with facts arrived at by scientific methods and evidence. This might be the grasping of an autobiographical genealogy in psychoanalysis, or the import of a critique of ideology. Foucault refers to an experience that “has the function of wrenching the subject from itself, of seeing to it that the subject is no longer itself …” (EW3, 241) It is an experience of apprehending a new network of significances, relationships and values that change our self-apprehension and the world in ways unforeseen. It is this sort of experience that has been covered over, invalidated by the demand for scientific method and evidence directed at objects independent of our mode of being. Foucault’s claim is not simply that psychoanalysis and Marxism are anomalous forms of knowledge lying somewhere between religion and science. As we will see, he argues that a historical event, the Cartesian moment, constituted the particular form of subjectivity which enabled the scientific worldview. By its compelling qualities of naturalness and giveness, this new regime of truth went on to occlude the older notion of truth.

CH. 4 also

4 GENEALOGY OF THE SUBJECT

Having set up the basic concepts underpinning Foucault’s final period of work, we can now trace his genealogy of the subject which enables him to locate both his own critical practice and science within modernity. I am persuaded by McGushin’s reading of this work as an ascesis, directed at the transformation of subjectivity, for both Foucault and the receptive reader (McGushin 2007). This ascesis is an urgent and indispensible task. “When today we see the … almost total absence of meaning, given to some nonetheless familiar expressions … ‘getting back to oneself’, ‘freeing
oneself’, ‘being oneself’, ‘being authentic’ etc. ...we may have to suspect that we find it impossible today to constitute an ethic of the self, even though it may be an urgent, fundamental and politically indispensible task, if it is true after all that there is no first or final point of resistance to political power other than in the relationship one has to oneself” (HS, 251-2).

This needs unpacking. The self is political since “almost meaningless expressions” (“freeing oneself”, “being oneself”, “being authentic” etc) obscure the real political question about how we should live, which cannot be decided by a scientific image of our “true” selves. We see, in Foucault’s analysis of the repressive hypothesis, the modern obsession with knowing the scientific nature of the self as a sort of prison. The host of proliferating discourses doesn’t liberate a “true” self but only one linked to further machinations of power (EW1, 271). The task is urgent because the modern regime of power/knowledge constitutes subjectivities which are necessarily already political. It seems impossible because, by the eclipse of the care of the self, the only truth of the self we have recourse to is the one presented by science. Induced normalisation produces widespread forgetfulness about the question of how to live a good life. Sciences such as economics tend to presuppose “true” goals of human life and society and so preclude, any serious consideration of such goals. Our modern attitude is dominated by an anxiety to find the truth by scientific inquiry into our natures, which will then tell us how to live. The subject transforms itself into a scientific object (EW1, 294). However, political resistance needs a new ethics.87

In outlining Foucault’s genealogy of the subject, I will draw on the lectures from 1981 to 1984, presented chronologically in terms of the genealogy, not Foucault’s biography. I will start with the 1984 lectures which show how, as a response to a crisis in democracy in ancient Greece, care of the self came to be linked to parrhesia or fearless truth-telling. After examining Foucault’s discussion on the Cynics I will then turn to what Foucault calls the “Cartesian moment”, when the ancient ascetic practices were left behind and, by a re-alignment of power/knowledge, modern sciences emerge. The chapter concludes with Foucault’s analyses of Kant, the Enlightenment and modernity, which frame his view of philosophy and science.

**Socrates, the Cynics and Parrhesia**

Parrhesia is a particular “mode of veridiction”, a way of truth-telling.88 Foucault had previously analysed discourses in terms of systems of discourses or power which enabled statements to be

---

87 “Recent liberation movements ... need an ethics but they cannot find any other ethics than an ethics founded on so-called scientific knowledge of what the self is, what desire is, what the unconscious is, and so on” (EW1, 255-6)

88 Parrhesia involves showing others what one is thinking by avoiding any kind of rhetoric, saying what is true because one believes it is true, and courage to speak this truth in the face of danger (Foucault 1983).
candidates for truth. Now he brings in the subject’s relation to herself and others. The mode of
veridiction is the way in which the subject recognises herself and is recognised by others as speaking
the truth (CT, 2-3). Foucault notes three modes of veridiction in the ancient world, that of the prophet,
the sage and the teacher-technician. Parrhesia, courageous truth-telling, is the fourth (CT, 15).

Foucault inscribes the history of philosophy within an overall perspective comprising two separate
strands that he traces from Plato’s dialogues, the Alcibiades and the Laches. In Alcibiades, parrhesia
proceeds in the form of an “ongoing ontology of the soul” requiring an account of the nature and
condition of the soul. In the Laches, Socrates sets up the parrhesiastic scene within which an
individual learns to care for himself, by focussing on his form of his life, the choices and practices
that make up his everyday existence. He learns to step back from himself, making himself into a
question, a problem and an object, not just of knowledge, but of care.

Both approaches, start on the basis of resisting the prevalent neglect of the soul by taking care of
oneself. However, from the Alcibiades, “the care of the self leads to the question of the truth and
specific being of that which one must be concerned about. What is this “me”, this “self”, we must care
about? … what do we discover in the mirror of the soul contemplating itself?” (CT, 246). This leads
to the Western metaphysical tradition which seeks to know the soul as an ontologically distinct reality
(CT, 160). Here we see the emerging pre-conditions for the study of man as an object that was to
come to fruition in the 19th century as the human sciences. From the Laches “… the care of self does
not lead to the question of what this being I must care for is in its reality and truth, but to the question
of what this care must be and what a life must be which claims to care about self” (CT, 246). The
Laches never raises the question of exactly what it is one must care for. Rather “… what is
designated as the object one must take care of is not the soul, it is life (bios), that is to say the way of
living” (CT, 127). This discourse, accentuating the living body which is subject to constant test, is
taken up by Cynicism.

The Cynics practiced a form of parrhesia which vehemently rejected the conventional social and
political world, aggressively challenging all values and practices. Cynicism obstinately tied the duty
to speak the truth to the true life. The conventional world was seen to be in opposition to the truth
which the Cynic modelled on nature, animality and natural instincts. The Cynic refused to have shame
in his body and its desires. He lived in public in near nakedness and poverty, haranguing those who
went by. In this uncompromising form of parrhesia, involving a life at odds with convention, the
bodily presence of the Cynic represents a permanent provocation to the society he finds false. The
Cynic makes his life a manifestation of truth. Established philosophical principles of reasoning were
displaced by deploying them, not in discourse, but directly in the Cynic’s way of life which confronts
people with the truth of their own lives in a manner irreducible to discourse. The true life appears as a
completely other life, incompatible with conventional life which it mocks. Incurring the anger of others, the Cynic risks his life by displaying it. This life is political, not leading to inner tranquillity, but militantly calling for another way of life. It is a form of care linked to unrestricted courageous truth-telling “which pushes its courage and boldness to the point that it becomes intolerable insolence” (CT, 165).

Whilst Christian asceticism became “the major medium of the Cynic mode of being”, from the 4th century, authorities entrusted the conduct of souls to pastors (CT, 181). The individual was no longer able to bring about her own salvation. A relationship to God was mediated by the requirement for obedience, leading to a mistrust of the self, which became “the object of an attentive, scrupulous, and suspicious vigilance” (CT, 334). By including the principle of obedience, Christian asceticism forged “a new relation of the self to the truth, a new type of power relation, and a different regime of truth” (CT, 320-1). Within this new regime, parrhesia was obscured and devalued (CT, 333). By the institution of the confession, the hermeneutic subject is constituted to exercise vigilance over his thoughts by examining and confessing them to a priest who interprets their true meaning.

Philosophy was gradually purged of radical Cynicism and stripped of spirituality, taming it to become a theoretical discipline, a body of knowledge more than a way of life. Despite the paucity of its textual legacy, Foucault sees Cynicism as a force in modernity, still posing the question of the true life as the other life. This legacy is not a doctrine, but “much more an attitude and way of being” (CT, 178). Cynicism still asks the “perpetually embarrassing question” about the philosophical life (CT, 234). Foucault argues that philosophy can’t be separated from a philosophical existence. It must “always be more or less life exercise.” Philosophy is a form of parrhesia - courageous, frank truth-telling - that constitutes its activity as a mode of subjectivity. The philosopher is bound to a way of life, a true life. This distinguishes philosophy from science which, by its objectivating stance, cancels out any subjective commitments except to propositional truth. Anyone can be a scientist if they have the skills and education. Philosophy however, entails a commitment that has consequences for a way of life.

This idea of a “philosophical life” seems peculiar to modern ears. It is best grasped through the work of Hadot, who regards ancient philosophy as a way of life, comprising ascetic exercises aimed at transforming the mode of subjectivity of the practitioner. Foucault argues that whilst we can still trace its subterranean path winding through philosophy’s history, this understanding is now almost completely lost to us. The philosophical question of a true life was taken over, at the end of Antiquity by religion, obscuring and covering over the ancient ethical notion of an aesthetics of existence. Later

---

89 The idea of bearing witness to the truth was found in spiritual movements such as the mendicant orders throughout the middle ages. “The Franciscans with their destitution, wandering, poverty and begging are up to a point, the Cynics of medieval Christianity” (CT, 182).
the institutionalisation of truth-telling practices in the form of normed, regulated, institutionalised science emerged, the result of a historical rupture beginning in the 16th century which narrowed the understanding of truth, invalidating what is outside the scientific method (CT, 235). The emphasis on deciphering and renouncing the self is eventually incorporated into the disciplinary structures of the human sciences (EW1, 274).

**The Cartesian Moment**

This historical rupture, the “Cartesian moment”, occurred in the context of tensions between the Church’s pastoral care of the self and the political art of governing populations. (EW3, 315; STP, 227-48; SD, 25). In the 16th century, “the art of government exploded” as a general social problem embracing government of oneself, the poor, families, armies, souls, children, cities and the state (EW3 201-2; PT, 44). A new form of political power arose which translated the spiritual government of souls into the political government of populations. It was into this context that Descartes inscribed his *Meditations* (STP, 230). The meticulous quality of Descartes method can only be understood if we bear in mind “that from which he wants to distinguish and separate himself, which is precisely these methods of spiritual exercise that were frequently practiced within Christianity and which derived from the spiritual exercises of Antiquity … ” (HS, 294).

Philosophy claimed back a form of *parrhesia* from Christian spiritual practices, such that “knowledge itself, and knowledge alone gives access to the truth” (HS, 17). The subject was no longer required to transform its being. There were conditions, but these weren’t spiritual conditions. The Cartesian moment leads to the government of life and populations becoming the concern of politics in the form of biopower backed by science. Descartes’ method becomes a new form of care of the self that ensures the proper conduct of the mind, thus giving the subject access to certain knowledge of everything that is useful for life.90 To grasp how this scientific worldview was opened up, and how the modern subject was constituted we need to examine Foucault’s analysis of Descartes’ *Meditations*. In his 1972 essay, Foucault offered a strikingly original interpretation that bears directly on the history of subjectivity and the fate of the philosophical notion of the care of the self 91 (HM, 550-74). Here we see an earlier Foucault, in the midst of his turn to genealogy, already concerned with the history of subjectivity and the self-constitution of subjects in their relation to truth.

The *Meditations* pose as a response to Descartes’ awareness of having previously accepted many false opinions as true and his lack of any firm criteria to distinguish between truth and falsity (Descartes

---

90 Instead of “the speculative knowledge taught in the Schools” Descartes new form of knowledge “is most useful in life” and will “make ourselves masters and possessors of nature.” (Descartes 1987, 78)

91 The paper was a response to Jacques Derrida’s criticism of Foucault’s interpretation of Descartes in HM.
1968, 95). His ultimate motivation is to gain a firm foundation on which to build certain mathematical and scientific knowledge. By systematically employing general doubt to doubt everything, Descartes aims to investigate what access the subject has to certain truth, and how the subject can be certain that this is the truth. Foucault sees Descartes as initiating a particular regime of truth, a way of binding the subject to truth. Descartes’ *Meditations* are not merely an investigation or an argument, but an *ascesis* that transforms the subject, leading him to a form of self-knowledge which opens up the scientific worldview which then replaces *ascesis* as an access to truth.

Foucault focuses on the passage in the first meditation in which Descartes questions whether he can actually doubt his senses. Certainly one can doubt what is “weak and distant”, but can one really doubt what is “vivid and near”? (HM, 554). A “practical syllogism” tells Descartes that he must doubt the senses because they have deceived him before (HM, 564). Yet Descartes simply can’t find a way to doubt that he “is sitting here next to the fire” even though reason tells him he ought. What is required is work on himself, to become a different sort of subject. Descartes’ hypothesis that the subject is dreaming is not merely another proposition in an argument. It modifies the subject (HM, 554-5). The dream hypothesis allows doubt to overcome “the actuality of the meditating subject (the place of his meditation, the action that he is carrying out, the sensations that strike him)” (HM, 565). This extension of doubt to the actuality of the meditating subject, “sitting here, next to the fire”, can only take place if the subject’s mode of presence to itself can be separated from this “system of actuality”. Doubt allows the subject to disclose itself in a way that is distinct from its actual being-in-the-world. By a form of distancing, a subject is posited who no longer needs to distinguish between the near and far, vivid and weak, but can view space and time neutrally from a disembodied scientific perspective. The dream seems as though it could be an actual state of the meditating subject, allowing one to free oneself from all representations (HM, 567).

What Descartes is doing in the *Meditations* is undertaking *ascesis*, modifying his subjectivity, as his readers are invited to do. But according to Foucault, this *ascesis* actually serves to establish a relationship between the philosopher and truth that will no longer rely on *ascesis*. After the 16th century, “we have a non-ascetic subject of knowledge. This change makes possible the institutionalisation of modern science” (EW1, 279). The modern philosopher doesn’t have to undertake spiritual exercises and transform herself. She just has to examine the evidence and follow the method and rules laid down for analysing parts. Philosophy can reclaim *parrhesia* and govern itself and others according to its own proper rules. Clear and distinct perceptions will enable philosophy and science to analyse things to see how they are composed. This is a new regime of truth, a new way of binding the subject to truth linked to a new application of power.
The Cartesian moment philosophically requalifies knowledge of self (*gnothi seauton*) and discredits care of self (*epimeleia heautou*) (HS, 14). Self-knowledge, in the form of the impossibility of doubting my existence as a subject, makes “know thyself” into a fundamental means of access to truth (HS, 14). The hermeneutic imperative of the pastoral, to know oneself continues, now guided by science rather than religion. As we have seen, this is increasingly generalised throughout our “singularly confessing society” in relationships between an individual and doctor, psychiatrist, psychologist, or psychoanalyst, the scientific expert who will interpret the truth of ourselves (CT, 5; HS1, 59). Instead of a notion of truth accessed through self-transformation, Cartesian knowledge “is knowledge (*connaissance*) of a domain of objects” (HS, 191). Rather than enlightenment, fulfilment, completion or salvation, the reward for this knowledge is indefinite scientific progress (HS, 19).

Whilst the new truth regime liberates subjects from the power apparatus associated with the Renaissance *episteme*, it also involves impoverishment by concealment. Discussing psychoanalysis and Marxism, Foucault suggests that the attempt to bolster these disciplines scientific status requires concealing their spiritual dimension and the relationship between truth and the subject (HS, 29-30).

Accompanying the Cartesian moment, political power is transformed by being no longer concerned with merely abstract subjects of juridical law, but subjects at the levels of individual bodies and the population in which it fosters life. Political power becomes concerned with people’s desires, the uses and conditions of their bodies, how they act and their relations with others. The life of individuals and populations is reduced to biology and public hygiene, which is forged into a scientific/technological support for the deployment of this rationality (SD, 182-3, 239-63). Biopower takes over the care of the self, putting forward biological life of the population as its ethical material to be normalised. Scientific knowledge of government, subjects and society is central to this project. Knowledge of the state must include wealth, resources and the population’s characteristics. Political reason tends to be reduced to the administration of life which is increasingly seen in biological and economic terms. As knowledge is regimented, philosophy becomes disqualified. It cannot constitute itself on the model of science. Nor can it delineate a distinctive field of study or set of problems.

5  **TRUTH, SUBJECTIVITY AND MODERNITY**

**Kant, Enlightenment and Critique**

In terms of the narrative thread set out in OT, if the Cartesian moment marked the shift from the Renaissance to Classical *episteme*, it is the figure of Kant that defines the shift to the modern. It is from this point in Foucault’s genealogy of the subject, that we see his own project in sharper focus. In the wake of the Cartesian moment, the practices of the self migrated to forms of government directed towards biopower - the biological and economic normalisation of individuals and living populations.
Philosophy was shorn of any residual affiliation to spirituality. If it was irrational to think that one could transform one’s relation to the truth, it isn’t surprising that Kant will discover human finitude in terms of necessity, rather than possibility. Kant gives the Cartesian moment a “supplementary twist … which consists in saying that … we cannot know the subject. Consequently, the idea of a certain spiritual transformation of the subject … is chimerical and paradoxical” (HS, 190). The Kantian subject is the condition of all possible experience and as such lies beyond experience itself. In Chapter 2, we saw Foucault analyse these paradoxical consequences, highlighting the oscillation between irreconcilable perspectives of transcendental subject and empirical object in Kant’s philosophy.

Foucault’s final period of work articulates the possibility of breaching limits and fundamentally transforming the subject’s mode of being. Finding valuable resources within Kant’s minor works, he takes up this theme in his 1978 lecture What is Critique, his 1983 and 1984 lecture series, and the 1983 essay What is Enlightenment? 92 To a point, Foucault offers a faithful exposition of this essay. Enlightenment is an ongoing task and an obligation that requires courage (GSO, 28). Since man is responsible for his state of immaturity, Enlightenment can only occur by “a change that he himself will bring about in himself” (PT 100). Foucault’s interpretation here draws on the ancient ethical qualities of the care of the self and the courage of parrhesia. What makes Kant’s text different is its engagement with the specific moment at which he is writing and because of which he is writing (PT, 104-5). Whilst philosophy paradigmatically dealt with eternal and universal truths, Kant’s essay asks about the present situation of the Enlightenment (PT, 98-9). What is philosophically distinctive about the present, is that it inquires into itself. This historical self-reflection is a “distinctive feature of philosophy as a discourse of modernity and on modernity” (GSO, 13).

Referring to the 16th century problematisation of government, Foucault links Kant’s account of the Enlightenment to “the critical attitude which appears in the Western world starting with what was historically, … the great process of society’s governmentalisation” (PT, 48). The question of government cannot be dissociated from the question of how not to be governed. Foucault defines critique as “the art of not being governed or better, the art of not being governed like that and at that cost… not being governed too much”, a formulation suggesting not ungovernability, but movement within a context imbued with governmentality. Critique seeks to expand the field of possible actions and thought in relation to a specific form of government.

Foucault’s understanding of critique and Kant’s account of the Enlightenment are both linked to resistance to power (PT, 45). On this basis, Foucault places himself within the Enlightenment tradition ensuing from Kant. However, Kant’s critiques are about knowing knowledge and its necessary and...
universal limits, whereas Foucault is concerned with problematising limits by revealing them to be contingent. We also saw how in OT, Foucault seemed uncompromisingly critical of Kant’s transcendentalism. On this basis, Habermas claims that Foucault has two radically different readings of Kant which map onto fundamental contradictions within his work. Referring to Foucault’s essay, Habermas asks “how such an affirmative understanding of modern philosophy … fits with Foucault’s unyielding critique of modernity” (Habermas 1992, 152).

To clarify this apparent discrepancy, we should see Foucault’s critique as part of a broader tradition of the critique of reason extending from Kant and Hegel through Nietzsche and Weber to the Frankfurt School and Foucault’s contemporaries. In this general sense, critique is any procedure which makes clear the conceptual pre-conditions of our thoughts, words, experiences and actions. This can be distinguished from more specific senses such as Kant’s transcendental and Foucault’s historical critiques. 93 Kant’s critique is transcendental in the sense that it is the type of inquiry that is not resolved by any ultimate appeal to sense experience. Experience takes place by virtue of the structure of our mental faculties which necessarily and universally constitute a priori conditions of objects in general. However, in Foucault’s historical critiques, the conditions of possibility are historical actualities. What has actually happened in history makes possible what actually happens in the present. Unlike Kant’s transcendental conditions, purportedly shared universally and necessitating particular forms of experience, Foucault’s historical actualities emerge in history to condition new frameworks of possibilities for thinking and acting in the present.

If Foucault thinks that Kant’s transcendental critiques can be located within a broader tradition of critique, he is entitled to claim to also be within this broader tradition, whilst rejecting its particular expression as transcendental critique. Foucault identifies his work with Kant’s essays on the Enlightenment, revolution, history and anthropology as types of critique quite distinct from transcendental critique. “This other critical tradition does not pose the question of the conditions of possibility of a true knowledge … but involves what could be called an ontology of the present, or present reality, an ontology of modernity, an ontology of ourselves” (GSO, 20-1). By “ontology”, Foucault doesn’t just mean being, but a way of being – a way that attends to the relations of power/knowledge by asking how and how not to be governed.

Foucault arrived at his historical form of critique by a Kantian move – asking after the limits and conditions of possibility of subjective experience. The fact that Foucault problematises these limits, rather than accepting them, is best understood as not a negation of Kant’s critical project, but a transformation, a radicalisation constituting a critique of critique. Whilst Kant thinks that to be

93 This useful distinction is from Koopman 2013, pp.109
autonomous, one must accept transcendentally revealed limits on reason, Foucault would see autonomy as contingently given by second nature, in other words, produced by discipline. As an object of disciplinary sciences, “Man” appears as a being who is able to both discover nature in himself, and manipulate it in accordance with what he sees as his potential. Here we see that the anthropological dualism discussed in OT - positive knowledge (discovered by the human sciences) and eschatological knowledge (which urges manipulation) - that is intrinsic to discipline. It is simultaneously positive knowledge of natural finitude and normative knowledge of a promise to be fulfilled. On one hand, a body is disciplined to make it obedient and useful. On the other hand, this same subject appears as free and capable of joining with other subjects in freely contracted relations. “The ‘Enlightenment’ which discovered the liberties, also invented the disciplines” (DP, 222).

Having made a distinction between Kant’s transcendental critique and the historical critique that Foucault calls for, we need to clarify this idea of critique in relation to philosophy. Foucault wants to draw a distinction between philosophy and the more radical critique of reason with which he aligns himself. Whilst philosophy might be considered the voice of reason, critique reveals reason’s underlying conditioned nature. Foucault’s ambivalence towards the modern philosophical tradition is seen by his location of critique at “the outer limits of philosophy, very close to it, up against it, at its expense …” (PT, 42). Critique only exists in relation to something other than itself - some institution, practice or discourse. Whilst close to philosophy it must maintain its distance to be critical. It is a means towards an undefined future, “an instrument, a means for a future or a truth that it will not know” (PT, 42). Critique eschews prescription. “Critique doesn’t have to be the premise of a deduction that concludes, “this then, is what needs to be done.” It should be an instrument for those who fight, those who resists and refuse what is” (EW3, 236).

Foucault’s critique suspends judgements. It is not negation, “a work of destruction, of refusal and denial, but rather an investigative work that consists in suspending as far as possible the normative system which one refers to in order to test and evaluate it.” (Foucault, cited in Lemke 2012, 61) Judgements operate to subsume particulars under already constituted categories, whereas critique inquires into the hidden constitution of the field of categories themselves. Foucault offers no normative grounding, rules or criteria to guide moral judgement. Rather than be told what to think and do, people have to craft their own ethics. Foucault thus challenges ideas of what critique is. He severs it from prescription, effectively challenging the idea that normative grounding is required for criticism. Foucault wants to see critical philosophy with more than the cognitive task of eradicating error. His critique doesn’t simply address lack of knowledge, but examines the limits that various truth regimes impose on autonomy and equality. It aims to bring to light aspects of the present ontological order which determine how one has to be in order to be, or to be recognised as, a certain way. This is clearly a separate endeavour from the dominant Enlightenment tradition of the “analytics
of truth” which inquires into formal conditions of truth, by looking for universal norms to separate rationality from irrationality (GSO, 20). In contrast, Foucault’s “history of truth” analyses historical conditions and limits of singular rationalities, relativising scientific rationality as one kind of rationality amongst a plurality. Rather than studying objects per se, his critiques investigate the “objectification of objectivities” (EW3, 238). Resisting assimilation into prevailing conceptual orders, they offer perspectives on those orders.

In its ongoing task of problematising the constitution of concepts, objects and subjects, critique strains to go beyond the limits that the tradition of philosophy has imposed on itself by experimental transgression towards an unspecifiable future philosophy. This experimentalism entails risk. To gain a critical distance from established authority, one must not only recognise the ways in which coercive effects of knowledge are at work in subject-formation, but risk one’s formation as a subject. Foucault suggests that critique involves something akin to virtue, as though virtue consists in risking established orders and stable identities (PT, 25). As a quality of ethics, virtue is not achieved merely by following objectively formulated or established norms but involves, more radically, a critical relation to those norms. Butler points to the “interrogatory” nature of critique which “presupposes a distance from all forms of morally motivated critique” (Butler 2001).

Critique is not just a critical attitude to power, but a self-transformative spiritual process directed at limits and the future. To question truth on its power effects and power on its discourses of truth requires a critical practice based on self-transformation. This “art of voluntary insubordination” insures “the desubjugation of the subject in the context of what we could call … the politics of truth”, or relations of power that by conceptually ordering the world, limit in advance what can count as true and meaningful (PT, 47). Just as Kant’s question about Enlightenment reflected on his present situation, so Foucault must reflect on what is different about his present, two hundred years later. He argues that modernity is not an epoch, but an “attitude”, an “ethos”, a “mode of relating to contemporary reality” (PT, 105). Rather than “a theory, a doctrine, or … a permanent body of knowledge that is accumulating”, this ethos entails “a philosophical life in which the critique of what we are is at one and the same time the historical analysis of the limits that are imposed on us and an experiment with the possibility of going beyond them.” (PT, 118).

As with his account of ancient ethics, Foucault expounds this modern ethos in terms of aesthetics, so distancing it from any foundationalist paradigm such as Kant’s transcendentalism. He wants an ethics based on freedom and self-determination achieved by the autonomous exercise of the will. Foucault highlights the relationship between ethics and aesthetic poesis. “This modernity does not ‘liberate man in his own being’; it compels him to face the task of producing himself” (PT, 108-9). Drawing on Baudelaire’s aesthetic modernity, he equates modernity with transgression and perpetual
transformation, the “feeling of novelty, of vertigo in the face of the passing moment” (PT, 106). In the struggle with modern power-knowledge-subjectivity, ethical self-fashioning is a force of resistance. An “aesthetics of existence” resists a “science of life” (Bernauer and Mahon 1994, 155). Thinking of existence in aesthetic terms releases it from the realm of scientific knowledge, and the endless self-decipherment and subjection to psychological norms. Foucault’s modernity accepts its condition as a contingent, localised, limited and ephemeral state of awareness, in contrast to Kant’s notions of timeless and universal reason. Rather than accepting Kant’s limits within which reason can operate, this new ethos must be “oriented toward the contemporary limits of the necessary, that is toward what is not or is no longer indispensable for the constitution of ourselves as autonomous subjects” (PT, 110).

Foucault wants to distance the Enlightenment ethos from humanism which encapsulates his aversion to any idea of a fixed human essence. Humanism is uncritical, borrowing images from religion, science or politics to substitute for the open-ended and undefined work of freedom. It can however, be opposed “by the principle of a permanent critique and a permanent creation of ourselves in our autonomy: that is, a principle that is at the heart of the historical consciousness that the Enlightenment has of itself” (PT, 112). By rejecting Kant’s transcendental critique but retaining the Enlightenment ethos, critique is no longer a determinate prescriptive relation to questions of knowledge, truth and falsity. It is now a matter of perpetual self-transformation, a process carried on within the understanding of the present, not from a universal understanding beyond time (PT, 108-9). This critical work of testing limits is what links ancient ethics to the philosophical ethos of the Enlightenment. The critical attitude towards the self reflects both an awareness of the contingency of one’s situation and a willingness to transform it.

Whilst retaining the Enlightenment ethos, Foucault challenges the Enlightenment’s determinate outcomes. He consistently criticised the modern myths of liberation: the progress of the human sciences, the liberation of the mad, prison reform, sexual liberation etc. In all of these he found, masked by Enlightenment reason, the forces of normalisation impoverishing and narrowing possibilities. The autonomy that underpinned Kant’s critique was legislative. One must conform to the moral law. But for Foucault, autonomy doesn’t involve freely binding oneself to a necessary and universal law. Rather, it involves the freedom to call into question all that is presented as necessary and universal. Whilst maintaining a link to the Enlightenment by an ethos, Foucault inverts the telos of Kantian critique. “If the Kantian question was that of knowing (savoir) what limits knowledge (connaissance) must renounce exceeding, it seems to me that the critical question today must be turned back into a positive one: In what is given to us as universal, necessary, obligatory, what place is occupied by whatever is singular, contingent, and the product of arbitrary constraints?” (PT, 113).

For Foucault, limits represent possibilities, not restrictions, beyond the known.
Philosophy and Science after Kant

Whilst Kant saw philosophy as an academic discipline alongside other disciplines, Foucault sees philosophy as a way of life. Whilst Kant’s project responded to what he saw as the scandals of metaphysics, Foucault responds to the main danger, the growth of modern power (PT, 54). This growing power is inseparable from the growth of scientific knowledge. Foucault sees one legacy of Kant’s epistemological critique as a stitching together of scientific positivism and the development of states into a science which plays “an increasingly determinant part in the development of productive forces and … state-type power increasingly exercised through refined techniques” (PT, 50-1). He is referring to the fact that, since the 19th century, science has been increasingly integrated with state power, which uses it to both develop productive forces and maintain and justify itself. This rationalisation has “effects of constraint and maybe of obscurcation, of the never radically contested but still all massive and ever-growing establishment of a vast technical and scientific system” (PT, 55). Whilst Kant’s critiques were blind to their own conditions of emergence within power/knowledge, they gave science confidence to operate on the basis of the necessary, universal finitude of the transcendental subject, within the secure limits of what could be known. Foucault wants to challenge the power effects that such confidence engendered, effects tied to the authority of science and “never radically contested”. We need to recall here that Foucault is not saying science is untrue. It is a question of what can count as true and meaningful. Since science emerged in the context of particular relations between the subject and truth, within particular constellations of power, it is also a question of forms of truth beyond those encapsulated by the methods of science.

Foucault locates himself in relation to critical theory by outlining the tensions arising between Kantian critical philosophy and the Enlightenment understood as ethos, a tension which led to increasingly sceptical questioning. “What excesses of power … is reason not itself historically responsible?” (PT, 51). In Germany, the Frankfurt School aimed to “show the connections between science’s naïve presumptions, on one hand, and forms of domination characteristic of contemporary society, on the other” (PT, 51). Foucault generally agrees with their approach, crediting Horkheimer as the first to highlight the fundamental problem: “The Enlightenment’s promise of attaining freedom through the exercise of reason has been turned upside down, resulting in a domination by reason itself, which increasingly usurps the place of freedom” (EW3, 273). He finds the Frankfurt School’s conception of the subject too permeated with Marxist humanism to accept that “what we need to do is not to recover our lost identity, or liberate our imprisoned nature, or discover our fundamental truth; rather it is to move toward something altogether different” (EW3, 275). Foucault doesn’t say what this “something altogether different” is. To spell this out would be to restrict the possibilities of

---

94 This is a travesty of the Frankfurt School position.
freedom by prescribing a notion of human nature, which is not something to be discovered but something made and hence able to be made differently.  

In France, the notion of the Enlightenment, with all the questions this involved, was less influential than in Germany (PT, 52). However, Foucault notes that with Bachelard and Canguilhem, the problem of the historicity of the sciences raised similar questions. “How is this rationality born? How is it formed from something which is totally different from it?” (PT, 54). Rather than the Frankfurt School’s analyses of the rationalisation of society or culture as a whole, Foucault’s background in the history and philosophy of science leads to analyses of rationalisation “in several fields, each with reference to a fundamental experience: madness, illness, death, crime, sexuality and so on.” Rather than invoke the progress of rationalisation in general, he analyses specific rationalities (EW3, 329).

Foucault thinks that since Kant, the question of the Enlightenment has been posed primarily in terms of knowledge, most often by asking “what false idea has knowledge gotten of itself and what excessive use has it exposed itself to, to what domination is it therefore linked” (PT 58-9). Foucault turns this around to ask about access “not to the problem of knowledge, but to that of power” (PT 59). He wants to locate the connections between power and knowledge. From the 1980’s the subject will be added, such that all three elements – knowledge, power and the subject – must always be analysed together. But in so doing, philosophical critique does not play the role of a science. Philosophy can’t “divide the true and the false in the domain of science” (GSO, 354). Nor can it say what should be done in politics. Nor should it free the subject. Rather, philosophy has to define the forms in which the relationship to self may possibly be transformed. Philosophical critique analyses the entwinement of power, knowledge and subjectivity, though from a location distanced from politics, science and the self.  

In other words, critique doesn’t propose solutions, discover anything new, legislate on truth or morality. In chapter 8, I will expand on the thought that whilst critique is not prescriptive, neither is it merely descriptive. It steps back, to make visible new possibilities in the relation of the subject to truth within contexts of power and resistance, thus inviting, rather than legislating or prescribing political action. “Where the role of science is to make known that which we do not see, the role of philosophy is to make seen that which we already see” (DE2 540-41, cited in Kelly 2009, 129). Whilst science discovers what we cannot see, philosophy allows us to see the same things in a light which reveals their constitution from our contingent practices tied to structures of power/knowledge/subjectivity.

95 “In the course of their history, men have never ceased to construct themselves, that is, to continually displace their subjectivity, to constitute themselves in an infinite, multiple series of different subjectivities that will never have an end and never bring us in the presence of something that would be “man” (EW3, 276).  

96 “Philosophy as ascesis, as critique, and as resistive exteriority to politics is the mode of being of modern philosophy. It was, at any rate, the mode of being of ancient philosophy” (GSO, 354).
In HS2, Foucault claims that the only worthwhile motivation for his work is a curiosity that doesn’t seek “to assimilate what it is proper for one to know, but that which enables one to get free of oneself” (HS2, 8). Philosophy is “the critical work” of “the endeavour to know how and to what extent it might be possible to think differently, instead of legitimating what is already known.” It explores what might be changed in its own thought, “through the practice of a knowledge that is foreign to it.” This foreign element provides “the assay or test by which, in the game of truth, one undergoes changes” (HS2, 9). This practice is a resistance to the modern philosophical neglect of the self which has now passed over into hands of doctors, psychiatrists, counsellors and scientific knowledge. This is what Foucault would see as the main danger today - a particular configuration of knowledge-power-subjectivity, a regime of truth bound up especially with the human sciences, to which subjects are subjugated. This configuration is itself the residue of past resistances, such as Descartes’ resistance against scholastic philosophy of the Renaissance *episteme*, or Kant’s resistance against the scandal of metaphysics. Yet when taken out of the context of the problematisations from which they emerged, these regimes of truth have become hypostatised and hegemonic forms of domination, calling for further resistance.

Foucault would insist there is no set of practices and discourses that can be considered final, true and universal. Referring to Habermas, Foucault argues that critique is not “a matter of identifying general principles of reality”, timeless universal principles, from which we can discern “what is true or false, founded or unfounded, real or illusory, scientific or ideological, legitimate or abusive” (PT 59, 60) There is no essence, no “natural” form of human life or sociality that science or philosophy can discover. Philosophical critique shows this, not by rejecting scientific truths, but by revealing the contingent constitution of constellations of power/knowledge/subjectivity as pre-conditions for the emergence of such truths.
Chapter 5
The Critique of Positivism to the Theory of Cognitive Interests

In Chapter 1, I discussed the value Habermas placed on unconstrained public discourse. Like Foucault, Habermas was, and has remained, politically engaged. As early as 1953, he invoked the “role of public critique as guardian” when criticising Heidegger’s silence after the war (Habermas 1953). During the 1960’s he used journalism to intervene in many and varied issues. He opposed aggressive campaigns against minorities, self-assertive patriotism and statism, the idea of a balance of terror in the face of nuclear destruction and a statute of limitation to Nazi war crimes (Muller-Dohm, 112-122). Habermas’ theoretical work reflected the same concerns for open and free discourse by responding to positivism, which he saw as a truncated form of reason stifling reflection. In many respects, Habermas shares Foucault’s critical stance towards reductive understandings of science which grant excessive authority to forms of reason that narrow human possibilities. However, he didn’t share the same suspicion of general theories or the ideals of Enlightenment humanism. Instead, Habermas sought to build a systematic basis to understand modern societies and advance conditions required for the maintenance of democratic will-formation. Rather than adopting Foucault’s “agnostic” stance towards scientific truths, Habermas integrated science into a philosophical framework to provide a view beyond the everyday experience of the lifeworld.

In this chapter, I will consider Habermas’ early critiques, drawing on essays in Theory and Practice (hereafter TP), The Positivist Dispute in German Sociology (hereafter PD) and Towards a Rational Society (hereafter TRS). These works coincided with a period of national reconstruction when the German state played a major role in the economy and when a “technocratic” state, in which major decisions would be made by scientific experts, seemed possible. Habermas argues against positivism’s understanding of science, and the technocentrism which it supports, as obstacles to egalitarian democratic will-formation. Whilst the confidence of technocentrism is today tempered by scepticism, Habermas’ critiques of this period remain relevant as an antidote to the ongoing temptations of scientistic thinking. I will then discuss Habermas’ substantial theory of rationality, Knowledge and Human Interest (hereafter KHI). Here he develops a differentiated account of reason that will grant a place to scientific rationality, whilst not identifying it with rationality per se. I will conclude by considering some criticisms which shaped Habermas’ next period of work.

97 Helmut Schelsky and Hans Freyer argued that technology has become autonomous. “Political norms and laws are replaced by objective exigencies of scientific-technical civilisation.” (Schelsky quoted by Habermas TRS, 59.)
1 THE POSITIVIST DISPUTE

Habermas’ Response to Positivism
Positivism can be traced back to the modernist project of the French Enlightenment and August Comte who charted the progress of knowledge through three stages, culminating in “positive” knowledge which eschews metaphysical abstractions, and appeals to invariable natural laws to explain observable phenomena and events.98 From the 1920’s, the logical positivists of the Vienna Circle, in common with earlier positivists, were committed to the theses of empiricism and the unity of science. They argued that natural science progresses by generating verifiable theories and any statement that can’t in principle be confirmed by empirical observation is metaphysical nonsense. Against this, Popper insisted that theories could never be confirmed but only falsified. The mark of a good theory was its falsifiability. Both logical positivists and Popper argued that all scientific enquiry, including the social sciences, should be based on empirical observations which can potentially verify (or for Popper, “corroborate”) statements, resulting in causal explanations and general laws.

Logical positivists held that the application of scientific methods to the study of social phenomena required rigorous avoidance of normative considerations, since value judgements did not admit truth or falsity. All science, including social sciences, had to be “value-free” (TP, 265). Scientific knowledge could be brought to bear on options, preconditions, or consequences, but the choice of ends itself was ultimately a question of values, requiring decisions, not facts. Including normative considerations in social inquiry could result in dogmatism and ideology, but not the cumulative progress of objective knowledge. Thus positivism sought to unmask normative aspects of arguments as “noncognitive”, “subjective”, “irrational”, so bringing to light what was seen as a confusion of facts and values. By revealing the pseudoscientific, ideological character of such arguments, positivist critique sought to establish continuity with the battle of the Enlightenment against ignorance, superstition and dogmatism.

Habermas’ response to positivism is seen in his contribution to an exchange initiated by Adorno and Popper, in which Habermas and Adorno defended social science against the reductive tendencies of positivism (PD). Habermas ventures beyond Adorno’s thinking, putting forward the idea of “a rationality with unconstrained validity in a domination-free discussion” (cited in Muller-Dohm, 113). Habermas was, at this time, concerned with the encroachment of functionalism and systems theory into social science. His critique is motivated by his conviction that positivism, blind to its own pre-

98 Kolakowski (1972) provides a good account of positivism.
suppositions, uncritically affirms the status quo. He argues that causal explanation is generally inappropriate to the social sciences because social processes typically involve processes of communication, which rest upon social actors making sense of their world by interpretations. By adopting the methods of the natural sciences as the model for all knowledge, positivists assume that the social scientist is distinct from the object of inquiry. Yet the positivist’s claim of value-freedom fails to account for the economic, historical and social contexts which condition the thoughts, perceptions and language in which observations are undertaken and articulated (PD 157-8). By abstracting from these background conditions, positivism restricts science to narrow interests in instrumental control and prediction, which it regards as the only access to reality.

Habermas thinks that the proper aim of social science is not instrumental control and prediction but critique and transformation. Popper’s rigid distinction between non-observable values and observable facts consigns critical evaluation of humanity’s true interests to the subjective status of decisionism, not susceptible to rational resolution and only resolved through arbitrary decisions (PD, 144; TP p. 265). The advantage of Adorno’s dialectical approach was that it “takes the social context of life as a totality which determines even the research itself” (cited in Muller-Dohm, 113). Scientific research, however objective, is still embedded in a lifeworld. However, the positivist’s analytic-empirical methods, assumed to be value-free, ensure that “simplified basic assumptions [are] chosen in such a way that permits the derivation of empirically meaningful law-like hypotheses” (PD 133). This false picture of society as subject to unchanging natural laws, reveals positivism’s ideological nature.

Habermas argues that scientific hypotheses are not tested directly by empirical observation. Rather the statement of a hypothesis is related to the statement describing empirical conditions (PD 152). But it isn’t clear what makes any given observation an instance of what is referred to in the statement. Positivists appeal to purely descriptive language stripped of vagueness. This presupposes a world directly imprinted and seamlessly transformed into protocol statements which form deductive steps leading to scientific knowledge. However, “there is no such thing as immediate knowledge” since “empirical data are interpretations within the framework of previous theories … they themselves share the latter’s hypothetical character” (PD 201-2). The procedure by which an observation is mapped onto a description can’t be formulated as a set of logically consistent self-contained rules. Popper, agreeing with Habermas on this point, argued that there must a discursive process within the scientific community similar to jury deliberation (PD 152). Picking up this thought, Habermas suggests there is no reason that compels acceptance of any given empirical observation as an instance of a statement of the test of a hypothesis. A decision must be made (PD 151). Yet according to the positivist, decisions relate to values and are relegated to the realm of the subjective and irrational. By positivism’s own criteria, “values are in principle beyond discussion” (TP 271). This restriction on reason renders positivism incapable of justifying its own interests. If a practical orientation is ultimately beyond
rational justification, then positivism’s commitment to empiricism is itself subjective and unjustifiable.

However, science is not arbitrary or irrational but open to rational determination by reference to a vast number of implicit beliefs that allow scientists to act and reach consensus. These implicit beliefs constitute the scientific lifeworld, within which science’s fundamental interest to control and predict “recedes into the background … disappear[ing] from the consciousness of those involved in the research process” (PD 155). This fundamental interest entails, as the sole admissible “value” (apart from methodological or logical values immanent to science) efficiency in the selection of purposive-rational means guaranteed by conditional predictions in the form of technical recommendations. This value cannot be seen as a value, “because it seems simply to coincide with rationality as such” (TP 264). By privileging this interest, positivism conceals a commitment to a particular form of rationality behind a veil of value-neutrality. Since science requires social consensus, there must be a form of rationality beyond instrumental rationality which is not mere rule-following but includes a communicative dimension for deciding what counts as an instance of a rule. “Argumentation differs from mere deduction by always subjecting the principles, according to which it proceeds, to discussion. … What can pass as criticism always has to be determined on the basis of criteria which are only found, elucidated and possibly revised again in the process of criticism” (PD 214).

Habermas notes that early Enlightenment philosophers saw the scientific study of nature as combating prejudice and unjustified authority of despots and the church (TP 257). Enlightenment reason was inherently practical with an interest in emancipation from natural and quasi-natural constraints. The decision to act was not external to reason (TP 254). But this relation of critical reason to enlightened practice was narrowed by positivism to one of prognosis and control. “As our civilisation becomes increasingly scientific ... science, technology and administration interlock in a circular process [in which] the relationship of theory to practice can now only assert itself as the purposive-rational application of techniques assured by empirical science” (TP 254). With positivism, the critique of values becomes universal and the possibility of self-reflection is abandoned. Assuming its own value-neutrality, science becomes blind to its inherent values of economy and efficiency, thus suppressing social practices not structured in these terms (TP 269). By refusing to recognise the historical emergence of social forms, positivism treats social relations in terms of ahistorical natural laws, applicable to the rational administration of society (TP 210). With the normative basis of society occluded, society and its members are treated as causally determined natural objects. Ultimately, the values by which we judge human actions risk erosion in a society understood in such narrow terms (TP, 273-4). By insisting on the extension of the methods of natural science to the study of society, positivism operates ideologically because this extension implies the naturalness of the status quo (PD,
Positivism fell out of favour in the 1970’s due not only to the critiques of Habermas and Adorno, but also the work of Wittgenstein, Foucault, Kuhn, Feyerabend and the growing field of history and sociology of science. Habermas’ critiques remain relevant for several reasons. The demise of positivism did not result in any clear consensus, but a plethora of approaches towards the philosophy of science. Within the social sciences, there remain fundamental questions about the relationship to natural sciences and appropriate methodological paradigms. Habermas’ overriding concern about positivism was as a form of ideology which suppressed and distorted thought and speech by disqualifying anything beyond the ambit of empirical science. This concern is sustained and developed throughout his career, including his later engagement with Luhmann’s systems theory and his more recent discussions of naturalism.99

The Technocracy Debate

Positivism found expression in “technocracy”, the idea of creating a “technical state” in which major decisions could be made by experts rather than politicians.100 This should be seen in the context of generally uncritical attitudes towards scientific and technological progress prevalent in the early 1960’s. During this period, naïve faith in science and technological automation led many to believe in the possibility of scientific rationality ushering in an age of prosperity in which scientific techniques could be extended to the administration of society. Habermas argued against these trends within the university, criticising the orientation of specialised disciplines to efficiency and utility. Universities should encourage reflection on the practical consequences of science for life, and on the relationship between science and society (Muller-Dohm, 120).

During the second half of 1960’s, the technocratic vision was opposed by increasing student activism, spurred on by geo-political issues such as the threats of nuclear arms, the Cuban missile crisis and the Vietnam war, leading to heated debates about the relationship of technology to society. As student protests grew, the authorities, encouraged by sections of the media, reacted repressively. Engaging with student concerns, Habermas argued for the democratisation of universities and a transformation of the capitalist structure of the economy along welfare state lines. Arguing against the violence of anti-authoritarian and anarchist groups, he saw no alternative to constitutional parliamentary democracy, despite the gap between the constitutional norms and reality. Habermas eventually rejected the protests as unreflective “action for action’s sake” (Specter, 90). The student bodies which had seen Habermas as a mentor, now distanced themselves or even opposed him (Muller-Dohm, 142).

99 See Habermas & Luhmann (1971); PDM, pp 368; TCA v.2; BNR, Ch 6.
100 Whilst this was discussed as a possibility in the 1960’s, it is clearly far from one today.
Habermas’ theoretical response to the technology debate was to highlight the relationship of scientific expertise and technology to political practice. Questions of life conduct require rational discussion not focussed exclusively on either technical means or the application of traditional norms. Habermas was, at this time, reading American pragmatism, in which he found a model for the critical interaction of the technical expert and the politician. The problem was that “new technical capacities, erupt[ing] without preparation into existing forms of life-activity and conduct”, were governed by “unreflected goals, rigidified value systems, and obsolete ideologies” (TRS 55, 60). These goals, value systems and ideologies must be made explicit and tested against, and adjusted to real possibilities. In this way, the ongoing development of technologies wouldn’t entail remaining bound to the same values. Value systems would be “partly confirmed, partly rejected, articulated and reformulated, or denuded of their ideologically transfigured and compelling character” by the challenges of technological possibilities (TRS 67).

One can think of examples such as reproductive technology or the vast range of possibilities made available by information technology to see how traditional values are transformed when confronted with previously unimagined possibilities. Given that technology (and scientific research to which it is tied) responds to our values which are themselves shaped by technology, Habermas advocates a relationship between technological possibilities and values, a “dialectic of potential and will” that could enlighten us about our tradition-bound interests in relation to what is technically feasible. He argues that this dialectic actually occurs, but without reflection and in accordance with interests for which no public justification is given or demanded. The challenge is to consciously set into movement a politically effective discussion that brings the social potential of technical knowledge into dialogue with values (TRS 61). Habermas advocates a “pragmatistic” model which promotes critical interaction between experts and politicians, involving a two-way communication rooted in social interests and in value-orientations of a given lifeworld (TRS 67). Here he restricts himself to considering the relationship between the politician (representing social interests) and the expert (on technical possibilities). We will see in Chapter 7 how the notion of critical interaction and feedback is expanded as a model of democratic society.

The tendency towards the ideological use of simplified, de-contextualised and overextended interpretations of science understandably provokes a demand for the positivistic separation of theory and practice. In this context, Habermas sees the notion of the “neutrality” of the sciences being reasonably employed against the short-circuiting of the connection between scientific expertise and a public vulnerable to manipulation by imbalances in social power and access to information. Nonetheless, as soon as this critique calls into question the communication between theory and practice as such, it “succumbs to the limitations of positivism and an ideology that makes science
impervious to self-reflection” (TRS 70). Habermas is not optimistic about the possibilities for an informed rational dialogue about science, although this is what he advocates. The ideals of rational and open debate have been undermined by the managerial activist state with its bureaucratised power, reinforced with specialised large-scale research organisations, to exclude the public as a political player (TRS 76).

Habermas insists that the challenge of controlling our species’ destiny in the face of unplanned sociocultural consequences of technological progress can’t be met by purely technological imperatives (TRS 60). We can’t assume that rationality extends from the capacity for technical control to the practical mastery of historical processes (TP 275). History cannot be made more rational by more technological control, however sophisticated, “but only by a higher stage of reflection, a consciousness of acting human beings moving forward in the direction of emancipation” (TP 275-6). Rather than simply presupposing technological solutions, this “higher state of reflection” would pose more fundamental questions about the goals of society which Habermas would insist cannot be decided in advance of democratic deliberation.

**Technocratic consciousness**

Even without the technocracy thesis, “technocratic consciousness” serves an ideological function by posing practical problems exclusively in terms of technological solutions. Habermas follows Marcuse in thinking that the key to analysing late capitalism is to understand scientific technology as a legitimating power, a role it assumed with the replacement of “free market” ideologies by state interventions designed to compensate for market failures (TRS 100-1). These interventions deal with social problems by attempting to eliminate dysfunctions and avoid risks, and hence require their interpretations as technical rather than as practical problems which require democratic discussion in terms of, for example, justice. By the fusion of science, technology, industry and state administration, “the development of the social system seems to be determined by the logic of technical progress … Scientific-technical progress “produces objective exigencies … which must be obeyed” (TRS 105).

Through the prism of “technocratic consciousness”, science and technology fulfil the ideological function of legitimating the exercise of political power over the heads of a depoliticised public. The basis of legitimation for a rational society thus becomes the growth of productive forces that issue from quasi-autonomous scientific and technological progress, with existing relations of production appearing as their technically necessary organisational form.

---

101 Habermas was writing well before the turn to neo-liberalism in the 1980's
As ideology, scientific-technological progress serves to impede making social and economic power relations objects of reflection (TRS 111-2). Technocratic consciousness differs from bourgeois ideologies which selected, amplified and distorted ethically loaded images of the good life and could thus still be traced back to a basic pattern of just interactions. By not projecting such images, “technocratic consciousness is less vulnerable to reflection than bourgeois ideology”, in which discrepancies between ideological projections and reality prompt reflection and critique. It renders inert interactions in the ordinary languages in which we are socialised and in which “domination and ideology … can be reflectively detected and broken down” (TRS112).

Habermas is not worried by scientific-technological development *per se*. He embraces modernity in which science has become the primary productive force. His concern is the orientation of practical life in accordance with technological imperatives in a society that addresses conflicts, not by discussion, but by de-politicisation. Foreshadowing his later colonisation thesis, subjects are both depoliticised and self-objectify in categories of purposive-rational action and adaptive behaviour, thus diminishing the capacity to maintain the inter-subjectivity of mutual understanding and communication without domination. Since these effects are concealed behind the expansion of the power of technical control, the new ideology requires a reflection that penetrates beyond particular class interests to disclose the fundamental interests of mankind as such. In KHI, we will see Habermas spell out these interests comprehensively.

**Habermas and Marcuse**

Habermas thinks that whilst science and technology are not politically neutral or “value-free” as the positivists claimed, they are nonetheless an expression of a fundamental *human*, not class, engagement with the world. He argues against Marcuse’s view of an alternative “new science” free from exploitation 102 (Marcuse, 1964). Marcuse thinks that there needn’t be a single technical rationality. Technologies are not means to independently chosen ends, but shape, and are shaped by, these ends. Science and technology are contingent and could be reconstructed to play different roles in different social systems. The current technological consciousness is constituted on the basis of society’s control by elites. Marcuse therefore proposes a revolutionary transformation of basic practices leading to a change in the very nature of instrumentality. Rather than treating nature as mere raw material to be exploited for power and profit, with the abolition of class society, human beings could adopt a more responsive attitude towards nature’s inherent processes and potentials.

Unconvinced, Habermas sees this as the same romantic rejection of modernity found in Horkheimer and Adorno (TRS 86). Their bleak vision presented instrumentality itself as domination (Horkheimer

---

102 Herbert Marcuse was initially a student of Heidegger who later joined the Frankfurt School with Horkheimer and Adorno. See Kellner (1984)
Technological control of objects violates their integrity and distorts the inner nature of the subject, leaving no alternatives but a retreat into art, nature or religion. Habermas thinks Marcuse makes the same mistake, jumping too quickly to the conclusion that the basic presuppositions of the natural sciences are exclusively “determined by class interest and historical situation” (TRS 85). Particular interests, including political interests, influence concrete historical constellations of technology, but underlying its various forms, Habermas discerns a basic logical structure relying on purposive-rational action as a feedback-controlled learning process based on a deep-seated anthropological interest. We can’t simply renounce technology or scientific research to which it is tied, since from the earliest tool-making it has played an essential role in the species’ survival. Technology is a general form of action reflecting a generic human interest in control, thus transcending particular political interests and politically neutral in itself. Marcuse’s alternative science isn’t possible because the grounding of science in instrumental reason is a precondition of the survival of the species as a whole (TRS 87).

On Habermas account, class interests don’t exhaustively determine specific historical forms of technology. Nor is the problem reason per se, but rather an imbalance between its two dimensions - the dominance of “labour” over “interaction”. By “labour”, Habermas is referring to the process of emancipation from nature, the growth of productive forces by the exploitation of technical knowledge. It is success-oriented, a form of purposive-rational action aimed at controlling the world. “Interaction” refers to social relations among individuals capable of communication and the pursuit of common understanding. Technocratic domination results from an imbalance which expands scientific technology to “the ‘historical totality’ of a lifeworld”, putting its particular forms beyond critique (TRS, 90).

Habermas misunderstands Marcuse’s utopian reflection on a new, emancipatory relation to nature. By referring to an “aesthetic” relationship with nature, Marcuse is suggesting a technology designed and applied with sensitivity to the inherent potentials of the medium and the context that it presupposes. This would involve certain resemblances to aesthetic practice in the sense of being exploratory and responsive rather than rule-bound and directed towards success and “conquest”. Marcuse’s point is that the dominant technological rationality favours the quantitative abstraction from qualitative, ethical and aesthetic values. Certainly, technical principles can be formulated in abstraction from any interest or ideology. However, as such they are mere abstractions. As soon as they are actualised, they take on a socially specific content relative to the historical subject that applies them. Feenberg offers the example of “efficiency” which, whilst formally no more than the ratio of inputs to outputs, is often understood in the unexamined context of workers’ resistance. When

---

103 Both Agger (1992) and Feenberg (1999) prompted my thinking by defending this claim.
applied to specific contexts, efficiency implies decisions as to what kinds of things serve as inputs and outputs, who can offer and acquire them and on what terms, and so forth (Feenberg 1999, 160). Technological design could *consciously* embody values, purposes and meanings rather than the abstraction and differentiation fetishized by industrial capitalism in search of profit. In comparison to the value-charged technical possibilities proposed by Marcuse, Habermas views technology as “neutral” since it is a project “of the human species as a whole, and not of an individual epoch, a specific class, or a surpassable situation” (TRS, 88).

Habermas insists that certain interests grounded in the very nature of what it is to be human are universal and have foundational importance. He cannot accept what he sees as a mixture of communicative and instrumental rationalities. Labour and interaction are generic categories, each with their own logic. While the specific historical forms of science and technology depend on institutional arrangements that are variable, their basic logical structures are grounded in the very nature of purposive-rational action. Like Marcuse, Habermas understands scientific knowledge as not the mere description of a universe of facts and their lawlike connections (KHI 304). Rather there is a “transcendental framework” of conditions that allow humans to cognise the world in the coherent way that they do. The problem with Marcuse, as Habermas reads him, is that he narrows this framework to particular historical and political interests, which he opposes to his alternative form of knowledge embracing natural and social sciences alike.

One difficulty with Habermas’ account is that his “logic” of labour remains essentially ahistorical. Borrowing from Gehlen, he suggests that technological developments can be modelled on the substitution of mechanical devices for human functions and faculties – a socially neutral sequence of prosthetic substitutions from digging stick to the computer (TRS, 87). But this doesn’t explain why substitutions are made differently in different social contexts. It is only by being culturally specific that one can explain concrete technologies, which don’t arise from “logic”, as a constant of human nature. Rather than a universalising approach, a historical approach, more like Foucault’s fine-grained analyses, would reveal the interaction of social values and technology. Rather than essences that mark out the invariant underlying logic of actions systems, Feenberg suggests that the history of science and technology is more characterised by family resemblances, and cannot be lumped together under a single category of human rationality (Feenberg 1991, 178). By equating science and technology with instrumental action governed by a technical interest, Habermas occludes the fact that technology is essentially developed or realised in the form of value and meaning-laden designs. Whereas science might arguably be understood in terms of a detached mode of observation and a reduction of experience to abstract mathematical relations, its social effects are only realised when actualised within technological designs.
Since their 1960’s debate, the influence of Habermas has overtaken that of Marcuse. As the question of technology receded from the political-cultural horizon, Habermas’ interest shifted to a communicative paradigm. He still distinguishes his own critique from that of the 1960’s radicals such as Marcuse and “post-modernists” such as Foucault, who he sees as “anti-modernists” for reasons I will discuss in Chapter 6. However in recent decades, radical criticism of technology has been reinvigorated by the environmental movement as well as analyses such as Foucault’s. I will return to Habermas treatment of science and technology in the conclusion to this chapter and again in the following chapter, to argue that Habermas has been unable to theorise the value-ladeness, variability and specificity of technology adequately to incorporate it into his critique of modern society.

The End of an Era

By the end of the 1960’s, the situation was rapidly changing. Rising counter-cultural movements, often embracing romantic anti-technology attitudes, voiced strong critiques of technocracy. With the publication of works such as Rachel Carson’s *Silent Spring*, the critique moved into mainstream society, which became more critical of technocentrism and, joined by scientists, demanded legislative controls (Carson 2002). With the revolution ushered in by figures such as Kuhn, Feyerabend and the historians and sociologists of science, positivism came to be seen as naïve and dogmatic.

It must be remembered that Habermas’ critique was undertaken during the ascendency of the social democratic activist state in which there was more confidence in the capacity of science and technology to make a better world. However his critique is not obsolete. Whilst scepticism towards science is now commonplace, there remains a stream of technology-inspired optimism. Advanced technologies are frequently proffered as obvious solutions to the complex problems - techno-agribusiness for resource distribution, futuristic engineering for global warming, or information technology for social crises. What these approaches have in common is not only their “top down” authoritarianism, but that ethical concepts, such as justice, fairness and responsibility are occluded, consequently avoiding any cost to those proposing such “solutions”. This suggests the ongoing ideological function of the conviction that practical problems are fundamentally technological.

During this period Habermas sought a middle ground between entrenched positions. Unlike positivism, he didn’t he see science as “pure” theory, free from contaminating interests. But neither did he accept Marcuse’s critique of the technological state, which reduced conventional science to class interests. Nor, like conservative such as Schelsky and Gehlen, did he embrace technocratic planning. Habermas’ early critiques are important in articulating distinctions that formed the basis of later theoretical developments. By adopting the categories of “labour” and “interaction”, Habermas’ foreshadowed the development of his differentiated account of rationality, firstly grounded in deep-seated anthropological interests, then in the normative structures of communication. We will see these
categories rearticulated as “system” and “lifeworld” in *The Theory of Communicative Action*, whilst “ideology” will be understood in terms of “systematically distorted communication.” Beyond such reformulations can be seen the same commitment to the analysis of the obstacles to egalitarian, democratic will-formation in modern societies.

2 KNOWLEDGE AND HUMAN INTERESTS

The concerns which motivated Habermas’ critique of positivism found their comprehensive theoretical articulation in KHI, an epistemological thesis of three different ways in which reality is known, each linked to a fundamental interest necessary for the survival of the species. His differential account grants separate places to natural and human sciences and philosophical critique. Habermas embraces the idea of the subject as inherently social, historical and cultural. His broadly Kantian stance sees the observing subject and observed object as not distinctly separate. He wants to salvage Kant’s claim of universal autonomy and responsibility in a way that accounts for the empirical processes of self-formation. Unlike Kant, who settled on the dominant forms of knowledge of his time, Habermas thinks that no particular manifestation of knowledge can stand as the prototype of knowledge per se. Knowledge is always in principle falsifiable (KHI 14).

Habermas wants to provide a systematic framework for critical social theory as a valid, distinct form of knowledge. Positivism’s dominance had undermined the tradition of epistemological inquiry in which natural science is just one category within a broader conception of knowledge (KHI 3-5). Kant’s core question of how reliable knowledge is possible had been corrupted into the narrow question of natural science and its methodology (KHI 4). However, by the end of the 1960’s, positivism’s notion of the unity of scientific method and universal laws was giving way to the relativistic implications of hermeneutics and various forms of historicism and culturalism. In the new front opening up between relativists and universalists, Habermas wants to defend a universalist approach to social inquiry, by radicalising Kant’s epistemology.

KHI aims to understand philosophy’s present situation by reconsidering the “abandoned stages of reflection”, located in the movement of philosophical thought from Kant to Marx (KHI, vii). In reconstructing this history, Habermas analyses the connection between knowledge and human interests to reflect the history of reason characterised by his earlier dichotomy between “labour” and “interaction”. This is now augmented by “domination”, to yield three distinct types of knowledge, each oriented by a deeper, non-subjective “transcendental interest”, from which we apprehend reality. These human interests - the technical, the practical and the emancipatory - are not transcendental in the Kantian sense. His project is empirical, employing reconstructive approaches developed in
developmental psychology and linguistics. Cognitive interests are “quasi-transcendental”, based in the natural history of the species and tied to the imperatives of the socio-cultural form of life, although not reducible to them. Labour and interaction provide viewpoints on the natural and social worlds, whilst critical philosophy, is specifically related to emancipation.

Because Kant’s transcendental subject lacks any evolutionary, developmental or historical constitution, Habermas draws on Hegel’s Jena lectures of 1803-6 to articulate his quasi-transcendental framework in terms of the categories of labour and interaction within which knowledge unfolds (KHI 15-16). However, Hegel failed to maintain this account of heterogeneous but interconnected patterns of self-formation (KHI 24). Habermas therefore turns to Marx’s articulation of the subject-object relationship, avoiding Hegel’s idealism by conceiving humans as actively transforming objects by social labour. Marx views the scientific-technological rationality built into labour as propelling the evolution of society, providing the impetus to transform rigidified ways of life. Whilst Marx’s critique of political economy transcends this narrow theoretical frame to include an ideal of social interaction, emancipated from systematic distortion by power, there is still a problem. According to Habermas, Marx tacitly conceives of reflection in terms of production, and so conflates natural and social science, leading to the understanding of economics in terms of natural laws 104 (KHI 44). A critical social science must avoid this conflation by distinguishing the rationalisation of purposive-rational action, or labour, from the rationalisation of communicative interaction which removes restrictions to self-reflection and communication. The productive forces don’t necessarily play a leading role in emancipatory movements. Their development can be a potential for liberation only if it does not replace rationalisation of the institutional framework which occurs by interaction. If interaction is neglected, rationalisation of productive forces will lead to an increase in technical control over objectified social and natural processes. Scientific-technological progress must therefore not bypass, or substitute for, the rationalisation in the medium of symbolic interaction which removes restrictions on communication. The Marxian distinction between the forces and relations of production must be replaced by the more abstract distinction between labour and interaction to enable critique. Habermas urges that we recognise and exercise a choice, not whether we utilise technological potentials, but about “what we want for the purpose of the pacification and gratification of existence”, a choice requiring “unrestricted communication about the goals of life activity and conduct, against which advanced capitalism ... puts up a strong resistance.” (TRS 119-20)

In developing his account of the three cognitive interests, Habermas draws on the insights of Peirce, Dilthey and Freud, the latter employed in reworking Marx’s historical materialism. Habermas insists

---

104 It is doubtful that Habermas provides a fair interpretation of Marx. See Grumley (1992)
that the sciences can’t be understood without reference to these knowledge-constitutive interests. Grounded in these fundamental interests, the viewpoints of three “sciences” – empirical-analytic, historical-hermeneutic and critical - disclose reality in different aspects. The first two viewpoints employ distinct categories such as “bodies in motion” or “acting and speaking individuals”. The use of such categories “implies an a priori relation to action to the extent that ‘observable bodies’ are simultaneously ‘instrumentally manipulable bodies’ whereas ‘understandable persons’ are simultaneously ‘participants in linguistically mediated interaction’” (Habermas 1975, 174). The third viewpoint takes up the already constituted objects of the first two interests in terms of emancipation. Habermas’ threefold differentiation allows autonomous areas of knowledge not reducible to natural science which is relativised as one form of knowledge amongst others. I will now consider KHI’s three “sciences”, before examining some criticisms and Habermas’ responses.

**The empirical-analytic sciences**

The survival of the species depends on successfully coping with threats and taking advantage of opportunities within the physical world. The technical interest which guides the empirical-analytic sciences addresses this dependency in the form of a learning process. Basic scientific statements reflect the success or failure of experimental operations within a fixed framework of possible experiences, given by the invariant relation of human beings towards their environment. Formulated into law-like hypotheses, statements enable prediction and control of nature. Empirical-analytic sciences are ultimately determined by the behavioural system of instrumental action which arises contingently in human evolution and binds our knowledge of nature with transcendental necessity to the technical interest in control over natural processes. Habermas is not making a psychological, or a sociological claim, nor reducing science to its utility. His analysis is directed towards the type of validity and the meaning of empirical scientific statements intrinsically related to purposive action. The control and prediction constitutive of empirical-analytic science are seen in its fundamental methodological and rational norms such as falsifiability, accuracy and replication.

Habermas turns to Peirce to articulate the distinction between labour and interaction within the context of natural science (KHI 91-139). Peirce’s pragmatism offers a form of transcendentalism that fully embraces fallibilism. For Peirce, a belief is that upon which we are willing to act. We orientate our behaviour according to our beliefs, which become settled as our actions succeed (KHI 121).

---

105 Habermas’ German “wissenschaften” has a broader meaning than “science” in English. The “empirical-analytic sciences”, refer to the natural sciences, but also include social sciences to the extent that they aim at producing nomological knowledge. The “historical-hermeneutic sciences” refer broadly to “the humanities” including the social, cultural, historical and human sciences to the extent that they aim at interpretive understanding of meanings. The “critical sciences” refer to as yet undeveloped methods of critical inquiry.
Knowledge is not grounded on immediate intuitive contact with the world or rational reflection, but advances through practical engagement, mediated by prior beliefs and practical expectations (KHI 97). New knowledge is generated in the context of a challenge to settled beliefs, the “irritation of doubt” as nature doesn’t respond as anticipated. If the behaviour that a belief gives rise to succeeds, the belief remains unproblematic. But when it fails, we seek to restore a settled state by supplementing, modifying or rejecting that belief. The capacity for rational-purposive control of nature involves an ongoing and cumulative learning process (KHI 120-1).

According to Habermas, instrumental feedback-controlled action, grounded in the category of “labour”, makes scientific knowledge possible, as a particular form of knowledge (KHI 120-6). Labour, our instrumental interaction with the world, enables the discernment of distinctions and the consequent synthesis of experience in terms of specific, contingently acquired, concepts (KHI 129-30). By instrumental action we establish which observable regularities are actually the causally necessary powers and properties which define objects. Since causal relations are not something directly sensed, but involve a counterfactual condition, this requires our active intervention. Causal necessity is not observed but is based on habit, or “fixity of belief”. Since no belief is certain in the Cartesian sense, “truth” is the property of a belief to withstand experimental tests in the long run. The experimental method allows truth to emerge when applied by an indefinite community of inquirers over an indefinite period of time. A belief is “true” if this ideal community of inquirers would accept it (KHI 91-5).

The settling of belief in this manner occurs at an ordinary everyday level. Taken-for-granted beliefs also enable scientists to interact within the scientific community in a manner not fully specifiable in terms of rules, but dependant on implicit cultural understandings embedded in the lifeworlds of scientists. If scientists are successful, beliefs about the right way to practice science are not open to question but remain part of this implicit background. It is only where science enters what Kuhn calls a “revolutionary” period that the paradigms on which scientific theories depend, are drawn out from the lifeworld to be challenged. (Kuhn 1996, 6-8) The difference between science and everyday belief formation is that science operates as an institution, a rule-governed disciplined structure that develops and refines pre-scientific knowledge by means that guarantee inter-subjective agreement, precise reference (e.g. standard terms and measurement conventions) abstraction (e.g. isolating the experimental situation from extraneous complexity) and systematising knowledge (universal assumptions linking theories). Isolated from mundane life processes, scientific feedback is reduced to a few significant forms, directed towards prediction and control. The cumulative progress of science is made systematic by research programmes increasingly planned in advance.
Inquiry can’t be justified by simply appealing to “reality” because if, as Peirce claims, all experience of reality is mediated by prior belief, and hence language, the accuracy of a hypothesis cannot be tested by an appeal to any immediate knowledge of the real (KHI, 99). With no access to anything independent of language, criteria for the progress of science are lacking. The way out for Peirce is to say that science assumes reality by its logic (KHI, 107). Habermas is sympathetic to Peirce’s commitment to the transcendental implications of this hypothesis of reality, which would serve to ground the intuitive notion that science is a privileged form of knowledge that progresses rather than merely changes. However with this hypothesis, a mere appeal to purposive-rational action meeting the world’s resistances to our beliefs isn’t sufficient to constitute knowledge. According to Habermas, Peirce is unable to respond to the challenge of Nietzsche who highlighted the historical contingency of beliefs and language, making the reliability of knowledge problematic (KHI, 118). To avoid perspectivalism which allows reality to be constituted in many ways, Habermas needs to ensure that the relationship between the interest in survival and knowledge is a transcendental one, rather than a merely empirical one.

For this relation to be transcendental, there must be a discursive process like the early Hegel’s notion of interaction. By such a process, the community of scientists determines whether or not empirical observations conform to a statement required to test the hypothesis, the process Popper saw to be like the disputation of a jury. Clearly, what our words mean, how we apply them, is not a matter for us to decide alone, but can only be determined inter-subjectively (KHI, 108). For Habermas, humans come to share an understanding of reality only to the extent that their engagement with it is mediated by concrete inter-subjective relationships to each other. Whilst the notion of inter-subjectivity is harnessed in KHI to explain the progress of science as increasingly adequate statements about the same reality, it is only in response to his critics that we will see Habermas clearly distinguish the constitution of experience from the inter-subjective constitution of knowledge.

The Historical-hermeneutic Sciences

The historical-hermeneutic sciences are guided by a practical interest anchored in the imperatives of socio-cultural life to maintain and expand the possibilities of mutual understanding. The inter-subjective and normative nature of this interest is not accessible to the analytical-empirical sciences (KHI, 176). Rather than an ontological distinction between different objects of inquiry, the same objects can be disclosed under different aspects. Human beings can be studied as objects of a science such as biology, or as communicative actors whose actions bear interpretation. Both forms of knowledge, though constituted differently, are achieved through the fallibilistic pragmatism that Peirce described in relation to the empirical-analytic sciences. “Both are set off by disturbances of routinised intercourse whether with nature or with other persons. Both aim at the elimination of doubt and the reestablishment of unproblematic modes of behaviour” (KHI, 175). In the case of empirical-
analytical science, doubt arises from the frustration of a feedback-controlled purposive-rational action. In the case of historical-hermeneutic sciences, doubt arises from the disturbance of a consensus, the non-agreement of reciprocal expectations of acting subjects.

Whilst the objects and laws of empirical-analytical science exist independently of the human mind, the historic-hermeneutic sciences are approached as products of human intentional action, of humanity coming to an understanding of its own created meanings. The empirical-analytical sciences bracket their own constitutive activity to apply hypotheses to alien material, whilst the historical-hermeneutic sciences grasp cultural events “from within”, since scientists are, after all, products of culture (KHI, 145). To elucidate this inter-subjective dimension of understanding, Habermas turns to Dilthey’s hermeneutics.

According to Dilthey, the empirical-analytic sciences, directed towards “explanation”, formulate particular experiences in a way that can be subsumed under universal laws. They “grasp reality with regard to technical control that, under specified conditions, is possible everywhere and at all times” (KHI, 195). The specificity and particularity of the individual experience is lost by abstracting from the richness of the context and smoothing out irregularities leaving only what is experimentally reproducible (KHI, 162). This is expressed in a “pure” language where context is irrelevant to interpretation and meanings are conventionally standardised. In contrast, the historical-hermeneutic sciences, directed towards “understanding”, invoke the particularity of experiences in ordinary language, by an inter-subjective process. They “grasp interpretations of reality with regard to possible intersubjectivity of action-orientating mutual understanding specific to a given hermeneutic starting point” (KHI, 195). Rather than subsuming under general laws, hermeneutics seeks to find expression for unique experiences, in all their rich particularity, within the generality of a shared ordinary language. Instead of a clear distinction between language and reality, as assumed by the empirical-analytic sciences, hermeneutic reality is constituted in language as the reality of meanings.

According to Habermas, Dilthey’s hermeneutics rested too heavily on the monological psychological concept of ‘empathy’, by which the experience of an historical agent could be re-experienced in the mind of the historian (KHI, 179-82).106 By seeking to ground hermeneutics on the possibility of an objectively correct interpretation, arrived at monologically, Dilthey failed to recognise the transcendental basis of the historical-hermeneutic sciences. By not adequately elucidating the inter-subjective basis of language and meaning, Dilthey positivistically reduces social and cultural

---

106 Habermas’ interpretation of Dilthey is disputed by a number of commentators, who consider he overplays the role of empathy and discounts development of Dilthey’s intersubjective ideas of expression. See Harrington (2001)
knowledge to natural scientific knowledge. Habermas argues that rather than attempting to experience the psychic state of the other, the opening of dialogue is required.

To consider such an approach we need to turn to the work of Habermas’ contemporary Gadamer who, Habermas argues, offers a more fully inter-subjective account (see LSS). Rather than founding knowledge on immediate insight, Gadamer sees all understanding as involving interpretation. For Gadamer, ongoing traditions are the medium in which “prejudices” or pre-judgements, as the conditions for understanding, are transmitted and developed. Rather than the reproduction of an original meaning, hermeneutical understanding involves the interplay of dialogue between tradition and the interpreter. The meaning of a text is inexhaustible in principle, since it must remain open to future interpretations in which the movement of history reveals new aspects which cast existing elements in a new light. This is important for Habermas because modern historicist awareness seems to break the continuity of tradition and claims that industrial society is no longer an ongoing unreflected extension of tradition, but rationally planned and technically mastered. According to such technocratic consciousness, a nomological social science could formulate general laws to which history could appeal as causal explanations of particular events. Habermas saw that by showing the dependence of interpretive understanding on an essentially historical socio-cultural prior situation of shared pre-understandings, Gadamer’s hermeneutics requires us to reflect on the relations of theory to history. Both Habermas and Gadamer recognised that the idea of society freed from historically conditioned consciousness (and a social science freed from the context of its interpretation) is illusory. The interpreter cannot assume a purely subject-object relation to his tradition since all his conceptual and normative orientations arise from this tradition. Since he belongs to the object domain being investigated, he can’t dominate it theoretically by the objectivating techniques of natural science.

Habermas agrees with Gadamer about this embeddedness of the social analyst in historically situated contexts. But he also sees the need for critical reflection to transcend this context. Habermas is concerned about Gadamer’s emphasis on the authority of tradition as medium of unreflectively shared pre-understandings. He sees Gadamer as elevating tradition to a position of unquestionable authority comprising the entire holistic background of pre-understood meaning that, by its very nature, cannot be made explicit all at once. Against this apparent conservatism, he argues that tradition is “profoundly transformed as a result of scientific reflection” (LSS, 168). When the everyday speech bequeathed by tradition is systematically distorted, it is only by reference to an “external” framework, that this comes into view. Habermas thinks that to identify hermeneutic inquiry with continuation of tradition is to emphasise participation and dialogue over distanciation and critique. Whilst he accepts the context-boundedness of all understanding, Habermas wants to mitigate Gadamer’s relativistic conclusions, arising from the situated character of understanding.
To achieve this, Habermas harnesses scientific theories of communicative development and social evolution as standpoints, external to the traditional lifeworld, which transcend the context of normal language competence. Such theories specifically reveal the empirical conditions of the formation and development of tradition. Whilst all social institutions depend on language as tradition, language depends on social processes involving power and domination which it does not articulate. If culture is viewed through the prism of an external perspective that brings into view its relation to the social, political and economic conditions of life, cultural tradition loses the appearance of self-sufficiency. Responding to Gadamer, Habermas notes that everyday linguistic understanding contains a possibility for reflection more critical than the tradition-bound, dialogical reflection that Gadamer has in mind. Certainly one can’t adopt the stance of a totally detached observer of one’s own language without at the same time leaving the realm of meaningful experience. Yet such radical distancing is not necessary in order to make a radical critique, since “every language is its own metalanguage” (Habermas 1980, 246). Here Habermas foreshadows the development of his context-transcending analysis of communicative action that we will discuss in the following chapter.

The debate between Habermas and Gadamer continued for a decade. They may have had more in common than either acknowledged. It is beyond the scope of this work to examine their misunderstandings or attempt a rapprochement. What is important is that Habermas’ arguments point ahead to the future development of his project. Habermas thinks that Gadamer devalues the methods of social analysis that go beyond normal linguistic competence. He will develop reconstructive theories of linguistic competence, cognitive and moral development, distorted communication and social evolution which can provide quasi-causal, though not nomological, explanations of social phenomena. Habermas will integrate this external perspective into his social theory. We see this beginning with his discussion of the critical sciences to which I now turn.

**The Critical Sciences.**

Thus far, we have considered Habermas’ epistemological thesis, yielding the analytical-empirical and historical-hermeneutic sciences, the former constituted by the technical interest underlying instrumental reason, the latter by the practical interest underlying communicative reason. However, the dependence of these sciences on interests embedded in empirical structures of human life suggests that reason is merely serving interests. Given Habermas’ critiques of both positivism and hermeneutics, neither form of reason is sufficiently rational. Habermas needs a critical perspective from a position outside empirically-bound interests. He thus draws on an emancipatory interest which guides a critically-orientated science. This interest directs the subject towards autonomous agency by engaging in self-reflection which recognises distorted and repressed content, “releasing the subject from dependence on hypostatised powers”. Guided by this interest, the “critical sciences”, or critical philosophy, can view the first two sciences to “determine when theoretical statements grasp invariant
regularities of social action as such and when they express ideologically frozen relations of
dependence that can in principle be transformed” (KHI, 310). As we will come to in my concluding
comments on KHI, Habermas subsequently acknowledges that the emancipatory interest he employs
actually underpins two distinctly different concepts of “reflection”, one of which is directed towards
conditions required for knowing speaking and acting as such and has no connection to emancipation.

The emancipatory interest constitutes reality in ways that second nature comes into view.107 It
recognises the determinate object as a product of the subject’s own activity, an activity contingently
shaped by powers otherwise unrecognised. Neither natural science, which only addresses first nature,
nor hermeneutics which assumes all products of human subjectivity to be accessible to natural
language, is sufficient for this task. The emancipatory interest requires self-reflection by a type of
reason, understood in terms of Kant’s sense of Enlightenment as freeing humanity from its self-
imposed tutelage by an orientation towards autonomy and maturity. At his 1965 inaugural lecture at
Frankfurt, Habermas linked this interest to communicative interaction, noting that its standards “are
exempted from the singular state in which those of all other cognitive processes require critical
evaluation. They possess theoretical certainty [since] the human interest in autonomy and
responsibility … can be apprehended a priori. What raises us out of nature is the only thing whose
nature we can know: language. Through its structure, autonomy and responsibility are posited for us.
Our first sentence expresses unequivocally the intention of a universal and unconstrained consensus”
(KHI, 314-15)

Habermas grounds the emancipatory interest in a developmental account of the species in which
communication becomes distorted and constrained by socialisation into traditions which serve to
conceal power underlying social relations. This conjectural scientific account provides the critical
sciences with a normative standpoint from which distorted communication can be recognised.
Freudian psychoanalysis and Marxist ideology critique are the paradigmatic examples of critical
“sciences” which aim for this sort of emancipation. However, neither are completely adequate due to
their incorporation of “scientistic misunderstandings”, specifically their self-understandings as natural
sciences. To rescue an emancipatory form of reason, Habermas must develop novel interpretations
that go beyond established disciplines. KHI should therefore be seen as laying out his preliminary
reflections as preparation to formulating a new conception of inquiry leading to critical reflection on
taken-for-granted beliefs, in the light of awareness of the formation of those beliefs within

107 First nature is what is immutably distinct from the subject, whilst second nature, or reification, is that which
initially appears to be a physical symptom but is revealed to be an intentional product of human agency, albeit
of an agency unrecognised as such.
constellations of social power (KHI, 310). To this end, Habermas develops non-reductivist aspects of Marx by incorporating Freudian themes into a revised historical materialism in order to unmask the illusions of second nature. For both Marx and Freud, false consciousness is a symptom of repression or domination preventing subjects from becoming autonomous.

Habermas treats psychoanalysis as a “depth hermeneutics” to be understood in terms of interpretation (KHI, 256-7). But rather than treating meaning as accidentally obscured, Habermas turns to Freud’s treatment of meaning as systematically distorted, thus calling for quasi-causal explanation, a task requiring a general theory of cultural evolution (KHI, 217, 257). As McCarthy points out, in his diagnosis of “communal neurosis”, Freud stepped outside the standards of given sociocultural frameworks employed in individual therapy, to take account of the cultural evolution of the human species (Freud 1985). Freud saw the motive for human society as the overcoming of scarcity, which varies historically according to the level of development and organisation of productive forces. The social institutions that inhibit the satisfaction of otherwise achievable desires can be viewed as communal neuroses, maintained by social power. In such an account, Freud is able to give interaction and language a more developed role than Marx who, according to Habermas, tends to reduce it to labour (McCarthy 1978, 84).

By incorporating Freudian themes into Marxist historical materialism, Habermas re-conceptualises notions such as power and ideology. Like individual neuroses, institutionalised power relations result in the rigid reproduction of behaviour. Unlike the open force of external compulsion, motives for action are censored from discussion and redirected towards substitute gratifications. Both in Freudian repression and Marxian reification, meaningful actions are seen by subjects as objectified causally determinate events. Because Freud treats dreams and neurotic symptoms as meaningful, Habermas can use them to develop a reading of psychoanalysis in terms of a theory of language. He looks to Alfred Lorenzer’s reconstruction of Freud which draws on developmental linguistics in order to decode the symbolism of neuroses (KHI, 256). What may initially appear as purely natural phenomena, such as memory loss or neurotic behaviour, is actually meaningful (KHI, 257, 228). A trauma or desire that is unacceptable for consciousness is made incomprehensible by sealing it within private language of the unconscious, such as dreams. Like Marx’s historical materialism, Lorenzer’s reconstruction of Freud provides a means to diagnose such disturbances and understand their deeper meanings. This understanding is distinguished from hermeneutic understanding by its explanatory potential. The meaning content can only be grasped by virtue of it being possible to explain the origin of the symptomatic scene by reference to the formative conditions in early childhood. The general theory offers an explanatory framework within which the patient’s fragmentary experiences may be reconnected.
Habermas brings to bear the idea of reason’s underlying emancipatory ideal by accounting for the overcoming of ideologies as the undermining of rigidified distorted communication and the power structures it legitimates. The telos of this ideal is “an organisation of social relations according to the principle that the validity of every norm of political consequence be made dependent on a consensus arrived at in communication free from domination” (KHI, 284). This counterfactual critical ideal, an anticipation of a future condition, intended to guide social transformation, is the core rational value that grounds human activity. This ideal, which will be developed in TCA, is the critical force that can challenge the hypostatisation of sciences as they are drawn into constellations of social power. The emancipatory interest grounds the transcendental status of all three knowledge-constitutive interests. Making the technical and practical interests explicit makes possible the reshaping of practices and interpretations. The technical interest enables the material conditions for political self-reflection whilst the practical interest provides traditions with stability (KHI, 210-1). Their justification is now directed towards political maturation, exceeding mere species survival.

According to Habermas’ notion of psychoanalysis, the “text” being interpreted is not initially recognised by the patient (KHI, 227). Within the therapeutic dialogue, a general theory provides a framework within which the fragmented accounts of the patient are analysed. The analyst offers interpretations, which only the patient can validate. The truth of emancipatory science lies in the patient who, by a process of self-reflection and dialogue which the analyst initiates, recognises a split-off portion of their life as their own (KHI, 257). However, Habermas was not suggesting that the critical theorist, in a privileged position, could extend psychoanalysis society-wide, to identify systematically distorted communication. Rather he was putting forward psychoanalysis as a model of the type of knowledge that includes an inherent will to emancipation. One should not view “the addressees of critical theory, still less society itself, as a macro-subject whose eyes the theorist aims to open; in a process of enlightenment there are only participants.” (Habermas 1994, 101, cf. TP 40)

Habermas’ critical science doesn’t exempt itself from belonging to the self-formative process on which it reflects. As his major work rebutting positivism, KHI argues against the idea of pure theory in the forms of either science or philosophy (KHI, 311). However, whilst the philosophical commitment to pure theory “succumbs to unacknowledged external conditions and becomes ideological”, within scientific practice this ideology has a protective function for the progress of scientific method (KHI, 314). “The glory of the sciences is their unswerving application of their methods without reflecting on knowledge-constitutive interests. From knowing not what they do methodologically, they are that much surer of their disciplines, that is of methodical progress within an unproblematic framework.” (KHI, 315). The problem is that scientism has colonised all reason. Whereas philosophy once marked out the legitimate place of science within a broader field of knowledge, “the philosophy of science that has emerged since the mid-nineteenth century as the heir
of the theory of knowledge is methodology pursued with a scientistic self-understanding of the sciences”. As we will see in Chapter 7, Habermas doesn’t think that philosophy, in this post-metaphysical age, can resume its role as adjudicator of other areas of knowledge. But neither can science take the role of a “first philosophy”, nor scientism insist “that we can no longer understand science as one form of possible knowledge, but rather must identify knowledge with science” (KHI, 4).

3 AFTER THE THEORY OF COGNITIVE INTERESTS

In KHI, we see Habermas articulate a position from which the sciences can be seen as partial forms of reason, not reason per se. His critique is not directed against the truth or falsity of the sciences, or scientific-technological development as such. It is directed against the threat posed by an increasingly dominant orientation of practical life towards the imperatives of science and technology which by adopting an ideological role, distort the communicative resources required for maintaining stable, free and open societies. In this sense, KHI is consistent with Habermas’ principal concerns I discussed in Chapter 1, concerns we will see throughout his work.

Given its ambitions, it is not surprising that KHI remains the book Habermas has been most ambivalent about, or that it attracted considerable critical discussion (Muller-Dohm, 136). Since Habermas “cannot claim more than the role of a prolegomenon” for KHI, his responses took the form of a series of adjustments, away from the transcendental analysis of constitutive interests, towards the communicative paradigm which will remain central throughout Habermas’s career. To frame the following discussion, I will characterise Habermas strategy as one of partitioning, whereby natural science is seen as only one form of reason, the limitations of which come into view from the perspective of critical emancipatory reason. Habermas wants to offer a diagnosis of social ills and some prescription, however formal and abstract, to address them. However, he can’t rely on norms passed down from tradition, but must turn to science itself, as the most authoritative form of knowledge, to provide an “external” perspective. Habermas’ critical stance thus harnesses conjectural but plausible developmental science which he employs not to merely describe change (as Foucault does) but, by establishing normative criteria, describe progress.

Habermas’ strategy stands in contrast to both positivism, which sought to reduce all science to a single method and hermeneutics, which sees all knowledge subject to interpretation. His theoretical construction brings into view, and grants legitimacy to, three fundamentally different and irreducible forms of knowledge. Two of these are bound to interests respectively grounded in biological and

---

108 Habermas responses can be found in TP preface, PKHI, Thompson & Held, Ch.12
social imperatives, but also linked to the third form of knowledge bound to an interest in emancipation. Whilst this reformulation of Dilthey’s distinction serves to combat any tendency to reduce all knowledge to natural science, this isn’t the only way to structure knowledge. As we have seen, Foucault’s “discursive formations”, cut across divisions, to include a vast web of heterogeneous discourses that interact to influence each other, developed to become, in Foucault’s genealogies, all-pervasive power/knowledge. Foucault, in Nietzschean spirit, would insist that Habermas’ partitions are more invented than discovered. Prompted by this provocation, I will consider a strand of criticism directed against Habermas’ theoretical structures.

It is not clear whether Habermas’ cognitive interests align with distinct areas of knowledge, or whether areas of knowledge actually overlap and merge in ways that render his division arbitrary. What is at stake is that Habermas’ analysis of natural science appears to mask their political character. The problem is not simply a matter of the subsequent political application of a form of knowledge that is “in itself” apolitical. Habermas insists that what is required to address the pathological aspects of modern reason is the re-establishment of a balance such that natural science is restricted to its appropriate domain. According to Habermas, natural science and technology are neutral because the technical interest is species-wide, rather than arising from particular social forms. The natural sciences and their associated technology share the characteristics of “labour”, guided by technical rules rather than social norms, with failure marked by inefficiency rather than social sanction. However, it is not clear that this characterisation of natural science holds. Whilst technical failure is different from social sanction, what actually counts as technical success or failure is socially determined, and enforced by sanctions internalised by actors who share their community’s values and goals. The anthropological interest in survival is not simply given, but relies on interpretations of fundamental distinctions such as harmless/harmful, useful/useless which are mediated by particular forms of cultural and social life.

Scientific research pursues certain questions, not others, employs certain categories, not others, and applies distinctions, not others, suggesting that by its very nature it cannot avoid responding to particular social values, rather than a neutral ahistorical technical interest enabling the species’ survival. Whilst as an ideal, Habermas’ technical interest may be universal, this universal interest never manifests as such, but only in the particular social and cultural contexts which give it form. Natural science only becomes manifest in contexts of speech acts within particular pragmatic contexts, or “translated” into technological projects, where other validity dimensions coalesce. Technology is realised in actual designs, unavoidably imbued with norms, values and meanings.

109 To take a simple example, a door is not a mere hole in a wall, a function whose efficiency seamlessly flows from the logic of the deep-seated interest to control and predict movement. It not only manifests this function, but signifies it by a “language” of socially sanctioned distinctions between servant and master, tradesman and guest, front and back, inner and outer private and public, etc. (which it reinforces or challenges).
In determining “efficiency”, it makes a crucial difference what we mean, in terms of the unquestioned contexts and values that frame the application of the concept. There are always alternatives, none of which are neutral, in satisfying the “universal” interests of species survival “efficiently”. For example, this interest might address climate change by the massive substitution of technical systems for natural processes. Alternatively, this same interest might be addressed by adapting ourselves to live more simply, foregoing the luxuries which a century ago were unthinkable but are now deemed essential for the “standard of living”. Efficiencies clearly relate to prior goals informed, consciously or unconsciously, by norms, values and meanings.

This complexity could come into view with a genealogy of the human and natural sciences, rather than an evolutionary perspective which reconstructs implicit human interests as the defining essences of branches of knowledge. Since the 19th century, shifts in the classification of knowledge have sharpened divisions between the sciences, whilst certain methods and forms of explanations remained linked (Daston & Vidal 2015). Daston and Vidal note, for example, that due to their interactive histories of shared epistemic values, practices and institutions, evolutionary biology and historical sociology may have more in common with each other, than either do with physics or demography (Daston & Vidal 2015). Rather than seeking fundamental anthropological interests as essences to define the unavoidable logics of distinct sciences, an historical analysis of actual science may yield a more complex picture which still supports a capacious conception of reason. Such a history would make explicit not only the strategic stakes in the positivist project, but also in natural science’s claim of neutrality.

In response to criticism of KHI, Habermas acknowledged that his theoretical construction conflated two senses of “reflection” in a way that weakens the interest in emancipation (Habermas 1975, 182-5). In one sense, reflection is critical insight into the unconsciously produced constraints and distortions to which subjects succumb in their formation. This insight, for which psychoanalysis is the model, in itself emancipates by revealing pseudo-objectivities. This emancipatory interest lacks the apparent primordiality of the other two interests. It is “derivative” because it “can only develop to … the extent that domination is institutionalised” (TP 22). Yet in another sense, “reflection” involves bringing something to consciousness that does not distort consciousness - the implicit anonymous universal rule systems which competent agents unconsciously follow, such as Chomsky’s deep grammar. This type of reflection (which Habermas calls “rational reconstruction”) is directed to the conditions required for knowing, speaking and acting subjects as such. Like Kant’s transcendental conditions for possible knowledge, these conditions not only constrain but enable, and their mere discovery is not emancipatory. “By learning logic or linguistics I acquire theoretical knowledge, but in general I do not thereby change my previous practice of reasoning or speaking” (TP 23).
Since rational reconstructions lack direct practical consequences, Habermas can no longer maintain the relation of reason to emancipation on which he wished to base critical theory (TP 23). There is now a gap between the “pure” disinterested reason of reconstruction and critique’s interest in emancipation. It looks as though critique is not reason itself, but a particular application of reason. In both historical materialism and psychoanalysis, a rational reconstruction of a general theory provides the background against which individual narratives are placed, so enabling the emancipatory movement, which is dependent on this prior reconstruction (TP 24). Rational reconstruction, independent of any ensuing emancipation, seems to be beyond the scope of the three interests. As a formal interest in theoretical enlightenment which does not have a concrete practical orientation, it is one of the conditions of possibility of critique and practical emancipation. But emancipation itself requires a further situated and engaged commitment.

Habermas also saw that he had conflated the constitution of knowledge with the constitution of experience which categorically divides the world into two domains corresponding to either the technical or practical interest. By not making clear the distinction between truth and the constitution of objects of experience, this appears to be a naturalistic reduction of the concept of truth to constituted experience. By tying knowledge too tightly to fundamental anthropological interests, it seems as though he leaves no room for a notion of truth that transcends particular contexts of experience. Without the discursive redemption of truth claims, we would have to conceive the progress of science as the production of new experiences rather than more adequate reinterpretations of the same experiences.

In response, Habermas introduced the distinction between the constitution of objects, relating to the categories of meaning by which we experience things (either as objects to be manipulated, or as persons, actions or utterances to be understood) and the constitution of knowledge which accommodates “the a priori of argumentative reasoning”, independently laying down the conditions of possible discourses” (Habermas 1975, 171). Empirical reality is thus constituted in the course of the pursuit of technical interests which underpin the survival of the species. But truth is removed from direct contact with experience and is asserted or denied only of statements in attempting to justify the speech acts of unreflective action (Habermas 1975 166). One difficulty with this distinction is that, given the theory-dependence of observation (whether perceptual or conceptual) it is not clear how the supposedly independent a prioris of experience and of argumentative reasoning interact in the complex ways that Foucault describes in terms of “seeing” and “saying” in BC (see Chapter 2). However, by making this distinction, Habermas can identify a form of communicative reason employed within discourse as common to all three areas of knowledge (Habermas 1975, 172). In the following chapter, we will see Habermas develop this form of reason as “formal pragmatics” in which
the “unity of reasoning” ensures that “in all sciences, argumentation is subject to the same conditions governing the discursive redemption of truth claims” (Habermas 1975, 172).

A further difficulty is that Habermas’ theory of cognitive interests appears to provide ultimate grounds, much like the totality and finality of first philosophy that he rejects. For example, he tells us that the fundamental interests “establish the viewpoints from which we can apprehend reality as such in any way whatsoever.” They are the “conditions of possible objectivity themselves” (KHI, 311). Habermas doesn’t want to assert the validity of a fundamental basis with the confidence of first philosophy. Yet he is also opposed to the sceptical notion that these concepts are merely arbitrary products of biological adaptation or convention. This generates a tension that runs throughout Habermas’ work which I shall return to in Chapter 8. Habermas understands that he can’t simply retreat from positive claims to knowledge. He wants to hold onto an emphatic concept of truth, and develop a critical philosophy to safeguard reason, freedom and justice. Only by retaining some comprehensive notion of reason can philosophy be saved from decisionism. Like Horkheimer and Adorno, he refuses to justify critique either as a privileged insight or a rigorous science. However he doesn’t think the distortions of reason are all-pervasive or that there aren’t standards to which we can appeal.

Thus far we have seen, in his early critiques of positivism, Habermas’ concern with defending a broad conception of reason. This was developed in KHI as an empirical epistemological thesis of quasi-transcendental interests which ground, and give legitimacy to, three distinct areas of knowledge. This ambitious project raised questions and difficulties regarding the distinctions and relations between the sciences, the notion of an emancipatory interest and the conflation of the constitution of experience with the constitution of knowledge. We will see how Habermas’ later work can avoid some of these difficulties by foregrounding communicative reason in terms of interaction, enabling the epistemological question of the constitution of objects to be detached from the subject, and retreat into the background. However, the shift from his theory of cognitive interests was not only motivated by these difficulties, but also by his awareness that the technical interest he had ascribed to empirical science was based on the positivist account of natural science which was increasingly looking less plausible. Since Kuhn, there had been a gradual weakening of the plausibility of the hypothetico-deductivist model, along with an increasing awareness of the role of interpretation and practical choice in scientific practice. Positivism was waning. In 1975, Habermas notes that “the criticism and self-criticism of scientistic assumptions are well underway” (Habermas 1975, 159). He notes the “confrontation of science theory with the history of science” within Anglo-Saxon traditions and mentions Kuhn, Feyerabend, Lakatos, Toulmin and Searle. Habermas will now shift from the critique of positivism towards a more comprehensive basis for his project. An epistemological approach will
give way to the communicative paradigm. We will also see new difficulties emerging in Habermas’ theoretical constructions.
Chapter 6

The Theory of Communicative Action

By the late 1960’s the differences between Habermas and sections of Frankfurt’s student body led to disruptions to seminars and lectures. With “gruelling and perpetual conflicts”, Frankfurt was no longer the stimulating environment it had been for Habermas (Muller-Dohm, 165). In 1971 he took up a position as a director of the Max Plank Institute at Starnberg. Remaining engaged with current issues, Habermas involved himself in shaping an institute that investigated the various aspects of the scientific and technical world. (Muller-Dohm 166] His magnum opus, The Theory of Communicative Action (hereafter TCA), was written at Starnberg and appeared as two volumes in 1981. Between KHI and TCA (1968 to 1981) Habermas developed a distinctive way of incorporating science into a normative and philosophical basis for social critique. As positivism’s influence waned, he framed this approach as an alternative to Niklas Luhmann's social systems theory which argued that modern societies have developed alternatives to normative integration, which can’t adequately deal with complexity.110

Habermas also distanced himself from the relativism of hermeneutics which sees understanding as inextricably historically and socially bound. To retain a critical perspective Habermas had to contain its relativistic implications. A theory of contemporary society which analyses symbolically pre-structured objects and events, by drawing on generalised empirical knowledge, would reduce the context dependency of understanding and leave room for critique. The paradigm of communicative interaction provided the perfect alternative to both systems theory and hermeneutics. The theme of communication, first raised in STPS, had re-emerged in Habermas’ critique of Gadamer’s hermeneutics to which Habermas contrasted the standpoint afforded by the theoretical reconstruction of communicative competence.

Throughout this period, Habermas’ fallibilist and empirical account of philosophy’s claim to universality and rationality, sought to push critical theory in a naturalistic, “postmetaphysical” direction by lessening its reliance on transcendental philosophy. The problem with the transcendental philosophy of the subject was that it obscured comprehension of the intrinsic intersubjective and dialogical nature of communicative action. Whilst avoiding transcendentalism, philosophy could still retain key universalistic claims, now reformulated in conjunction with human sciences to enable an

---

110 Systems theory had been developed in the mid-20th century by Talcott Parsons and was revitalised in Germany in 1970’s by Luhmann. As a preliminary definition, systems theory is an analysis of a series of interacting parts that are at least partly self-regulating, much like organs of an organism.
analysis of the general conditions of rationality manifested in various human capacities. Habermas’ formulation of a “language-theoretic foundation” to social theory brought his earlier interest in communicative reason to the fore, filling out the claim of his 1965 inaugural lecture (TCA1 xxxix): “Our first sentence expresses unequivocally the intention of universal and unconstrained consensus” (KHI 314). The theoretical reconstruction of the everyday communicative competences of social actors will form the foundation of Habermas’ theories of society, rationality and modernity.

In developing this theoretical basis, Habermas drew on a number of distinct theoretical strands - general theories of communication, systems theory, developmental psychology and social evolution. He employs the speech act theory of Austin and Searle to formulate a general theory of communication, reconstructing the universal pragmatic pre-suppositions of speech and action. This “formal pragmatics” aims to render theoretically explicit the pre-theoretical know-how underpinning basic communicative competencies. The rational reconstruction of formal competences enabling communicative acts sustaining everyday social life as stable, meaningful and ordered, explains the formation and ongoing reproduction of society. Uncovering implicit universal rules, common to language as such, it anticipates an ideally rational form of communication directed towards a counterfactual uncoerced consensus (Habermas 1980, 267). Formal pragmatics, signalled in his 1970 Christian Gauss lectures at Princeton, became the basis of Habermas’ linguistic turn in social theory (OPSI, 3-104). He further develops the theory of universal communicative competence with the pivotal 1976 essay *What is Universal Pragmatics?*, giving it definitive form in TCA. This paradigm of communicative interaction preserves the aspirations of the critical theory tradition, suitably modified for contemporary conditions and needs.

Whilst Habermas remained concerned to dispel any positivistic notions of cybernetically self-regulated societies, he also recognised the possibilities of systems theory. In *Legitimation Crisis* (hereafter LC), he had distinguished the lifeworld as a “sociocultural subsystem”, co-ordinated only by natural language and resistant to functional specification, due to its holistic and normative nature. In TCA, he draws a firmer distinction between system and lifeworld, developing the lifeworld, not as a subsystem, but as “a concept complementary to that of communicative action” (TCA2, 119). As we will see, this distinction marks off certain areas where social systems theory has priority, whilst keeping it clear of other areas. Adapting aspects of Parsons’ and Luhmann’s theories, Habermas brings into view objective interconnections beyond those subjectively intended or expressly articulated. He employs systems theory to gain a perspective beyond that of interpretive sociology, which is incapable of seeing where “language is also a medium of domination and social power” (LSS, 172).
Borrowing from Piaget and Kohlberg, Habermas radically recasts historical materialism in terms of the developmental stages of societies (CES, 95). Whilst developmental psychology provides a general theory of socialisation in terms of the acquisition of communicative competences, Weber, Durkheim and Mead provide theories of social evolution enabling a systematic reconstruction of historical situations. By employing these speculative scientific theories, Habermas can view the development of society as a learning process comprising a series of stages, each of which depends on the previous stage. He is thus able to establish provisional standards by which society can be analysed, so retaining the normative standpoint essential to critical theory.

By incorporating the scientific strands of formal pragmatics, systems theory, developmental psychology and social evolution, TCA provides a comprehensive theory of modern society, critically aimed at forms of thought that reduce the possibilities of freedom. As well as the ideological identification of scientific reason with reason per se, the relativism entailed by some of the reactions against such positivism - Gadamer’s hermeneutics, the Wittgenstein-inspired sociology of Winch, Heidegger’s historicism - must also be avoided. Habermas counters such relativistic tendencies by drawing on science, recognising that whilst scientific authority can fulfil ideological functions, it can also demand discursive justifications and critique arbitrary structures of prejudice (LC, 84). Acutely aware of science’s capacity to both liberate from prejudices and entrench new prejudices, Habermas is not exempt from the hazards entailed by this ambiguity. In this chapter I will firstly discuss Habermas major theoretical components - formal pragmatics, the system/lifeworld distinction and social evolution. (Sections 1-3). Having setup this framework, I will critically assess its implications in terms of modernity and its pathologies, particularly in relation to science and technology. (Sections 4-8)

1 FORMAL PRAGMATICS

The basic idea of formal pragmatics is that speaking subjects know how to achieve goals, produce things and perform acts without relying on an explicit account of rules, criteria or schemata. Habermas distinguishes three functions of language (TJ, 52). None of these functions – the representational, expressive and communicative - articulated by the traditions of formal semantics, hermeneutics and pragmatism gives a full account of linguistic functionality on its own. Habermas aims to unify these three dimensions of language in his formal pragmatics, as the competences humans need in order to communicate. Drawing on Austin and Searle, he is concerned with the illocutionary component, that aspect of language that Austin calls “doing things with words”.

111 McCarthy suggests that Habermas is “seduced” by the illusion of a rigorous science. (McCarthy 1991, 180)
entailing the establishment and reproduction of specific relationships constitutive of everyday social life (Austin, 1962, Searle 1969). One simultaneously addresses matters at hand by cognitive utterances with propositional content and, at the same time, employs speech acts with illocutionary content to engage in social relations. Promises, for example, are explicitly illocutionary, but also presuppose propositional content, which can be challenged, quite apart from the relationship established (CES, pp34).

Speech acts are only successful if they have the force to generate freely-entered interpersonal relationships between subjects. The illocutionary force (i.e. the speaker's particular intention in producing that utterance, for example, asserting, questioning, promising, advising, etc) derives from the recognisable and sincere willingness of the speaker to enter into a particular social relation, to accept its obligations and draw the relevant consequences for action (OPC, 83-4). The role of communication in the establishment and reproduction of specific relationships constitutive of everyday social life is assured by the “binding and bonding force of the speech act” aimed at facilitating voluntary cooperation by raising, implicitly or explicitly, three claims to validity - the truth, rightness and sincerity of a speaker’s utterance. Only one of these claims is thematised, depending on whether the speech act is constative, regulative or an avowal. The other claims recede into the pre-thematic holistic background of the lifeworld.

Each of the three validity claims corresponds to one of three “formal world concepts” – the objective, inter-subjective and subjective. Any of these validity claims can be drawn out of any utterance and subjected to the same inter-subjective norms of redemption. Meaning is always framed in terms of these three presupposed claims applying to language as such. They are transcendental in a weak sense of providing the conditions of possibility of communication aimed at mutual understanding, not the necessary conditions for possibility of knowledge or experience as such.

Communicative action, directed towards unconstrained agreement in language as the fundamental medium for social co-operation, involves speech-act immanent commitments. Constative speech acts (such as reporting, asserting or narrating) which mark the distinction between appearance and reality by thematising the validity claim of truth, entail “an obligation to provide grounds [and] … contain the offer to recur if necessary to the experiential source from which the speaker draws the certainty that his statement is true.” Regulative speech acts (such as requests, commands and warnings) which mark the distinction between what is and what ought to be, by thematising the validity claim of rightness, entail “an obligation to justify, … to indicate if necessary the normative context which gives the speaker the conviction that his utterance is right.” Expressive speech acts (such as revealing, admitting or deceiving) which mark the distinction between the authentic self and the expressions in which it appears, by thematising the validity claim of truthfulness, entail “an obligation to confirm, to
show in the consequences of his action that he has expressed just that intention which actually guides his behaviour." Thus the basis of the illocutionary force that leads to the hearer understanding and freely accepting the speaker’s intention is rational because “speech act obligations are connected to cognitively testable validity claims” (OPC, 85).

The communicative competence that Habermas wants to explicate is most developed in the later Wittgenstein’s idea of a language-game (OPST, 54). A competent player of a language game can follow a rule into novel situations without recalling or describing that rule in any detail. Communicative rules are not simply rules of correspondence between words and things, but rules about the way we use language, its appropriateness within particular social contexts. (OPST, 73) Such rules are inter-subjective, requiring others to hold us to them, as we hold ourselves and others to them. Wittgenstein conceived language as a multiplicity of ad hoc games in which the rules can vary according to the cultural context. Habermas is looking for universal rules that underlie all language use, independent of the user’s particular language games or culture.112

Habermas’ account is not simply a description of the mechanisms by which competent agents establish and maintain relationships. Since the illocutionary force has a rational basis, it is also normative, displaying not only how relationships are conducted but how they should be conducted. It is oriented towards the goal of critical theory – a life free from domination – anticipated in every act of communication. Habermas’ entire project rests on this possibility of providing an account of communication that is both theoretical and normative. Formal pragmatics is a continuation of his ambition to chart a comprehensive notion of reason as a viable alternative to instrumental reason. It yields knowledge that is empirical rather than a priori and fallible rather than certain, although without being reducible to an empirical science which causally explains language usage. However, it is directed to invariant structures and conditions and raises universal, but defeasible claims. Whilst it explains the implicit know-how to generate speech acts, the rules that form the conditions are not innate, but acquired by socialisation. Like hermeneutics, it relies on tacit understandings and normative expectations about speaking, acting and knowing. These are unavoidable – we can’t imagine substituting another set of norms.

When misunderstanding persists, an option is to continue seeking mutual agreement by “discourse”. This involves a suspension of the assertive force of an argument’s propositional content and the explicit thematisation of its validity claims as hypothetical. Discourse is guided by normative considerations much like those Habermas identified in STPS as residual in the bourgeois public

112 In Habermas’ view, without a notion of communicative competence underlying all language use, Wittgenstein lacks critical purchase and slips into perspectivalism (LSS 148).
sphere - the authority of the better argument should assert itself against status, it is possible to challenge areas of concern previously monopolised by authorities and the public is in principle inclusive (STPS, 36, see also Chapter 1). These communicative norms, directed towards freely achieved consensus, serve as a critical resource by which we recognise illegitimate forms of social organisation. Legitimacy is not merely a matter of the current acceptability of norms, but rather the way in which the norms have been achieved, which can be judged by the “ideal speech situation” (CES, 188). This standard is not intended to delineate some ideal communication or community, but to specify what is counterfactually presupposed in every communicative interaction. Like Kantian “regulative ideas”, it functions as a critical reference point, critical because validity is not defined in terms of particular contexts, but transcends all contexts.

In explicating this standard, Habermas tempers Adorno’s worry that imbalances of power always covertly invade language and thought. He pragmatically accepts the taken-for-granted world as a starting point, from which competent agents make a number of assumptions, almost invariably counterfactual. Agents assume that communication will not be distorted by imbalances of power, that all participants are equally able and willing to redeem the validity claims implicit in their utterances and that participants in communication relate to each other as responsible subjects (OPSI, 148). Social agents assume these counterfactual “fictions” as the critical standard inherent in communicative action (OPSI, 102). But there can be good reasons to abandon these assumptions. The ideal speech situation can be disturbed by strategic action, not directed at reaching mutual understanding. Here, a speaker conceals their intentions to act according to instrumental prediction of the other’s responses. He raises claims not because he can justify them, but rather to get “the addressee to draw his conclusions from what the speaker indirectly gives him to understand” (OPC, 332).

There is a further class of deviations from the ideal speech situation, which agents can’t recognise. In “systematically distorted communication”, actors are constrained by covert power below the level of conscious articulation. Habermas’ earlier account in KHI had drawn on psychoanalysis, whereby defence mechanisms inhibit the public articulation of neuroses, which are expressed in the “split off” symbols of a private language. From the perspective of formal pragmatics however, systematically distorted communication shows itself as the inability to raise validity claims to intelligibility or truthfulness. This is not due to the lack of hermeneutic efforts, but rather reflects the inability to deal with an external situation, say unjust power relations, which are manifest in the disruption to communicative competence (OPSI, 159-64).
2 LIFEWORLD AND SYSTEM

From the early 1970’s, Habermas developed formal pragmatics as a means to incorporate communicative inter-subjectivity into the form of a critical social theory. However, a substantive sociological theory remained undeveloped. TCA addressed this gap by developing a theoretical basis binding formal pragmatics into a comprehensive critical social theory. A key aspect is the elaboration of the concepts of lifeworld and system to offer an account of modernity in terms of its social, linguistic and intersubjective nature. Habermas maintains that both concepts, lifeworld and system, are essential to social theory. His analysis reveals pathologies resulting from the distortion of the lifeworld’s communicative resources by systemic colonisation of the lifeworld.

In TCA, Habermas steers a path between the relativism of Gadamer’s hermeneutics and the objectivism of Luhmann’s systems theory, whilst maintaining the perspectives of both participant in, and observer of, social practices. Although in LC, Habermas had employed systems theory within his analytical framework, he remained critical of Luhmann, for marginalising the subject’s active constitution of social worlds (LC, 6).113 At this stage, Habermas had not developed the concept of the lifeworld beyond its origin in Husserl’s phenomenology as "a naive familiarity with an unproblematically given background" comprising the stock of unthematised meanings, values and assumptions which individuals within a society share (citing Schutz and Luckmann in TCA2, 130). Now in TCA, he enriches the lifeworld concept to fully accommodate its inter-subjective linguistic nature, to serve as a counterpoint to the concept of system and account for the reproducibility of social relations.

The Lifeworld

Habermas’ develops the notion of lifeworld as a resource that harmonises intended action orientations by normatively securing communicatively achieved agreements. Communicative action, oriented towards mutual un-coerced agreement, is "encompassed within the horizons of the lifeworld, however blurred these may be ...” The lifeworld now becomes its condition of possibility. “The lifeworld is the intuitively present, familiar and transparent vast and incalculable web of presuppositions that have to be satisfied if an actual utterance is to be at all meaningful, that is valid or invalid" (TCA2, 131).

Communicative action proceeds by ongoing redefinitions of "situations" or “segment[s] of the lifeworld context of relevance that is thrown into relief … and articulated through goals and plans of action” (TCA2, 122-3). Claims initially in the background are made explicit, put on the table to be clarified, accepted or rejected. Unlike Husserl’s notion, Habermas’ lifeworld is intrinsically inter-

113 The debate between Habermas and Luhmann can be found in German in Habermas & Luhmann (1971). Some helpful English comments are found in McCarthy (1991, pp 155)
subjective, enabling understanding to be reached and inter-subjective breakdowns to be repaired by judging thematised claims on the basis of unproblematic background convictions. Situations are ordered by validity claims allocated to one of the three formal world concepts which emerged by their differentiation during the lifeworld’s evolution. The socially shared lifeworld has a quality of certainty, as prior to any possible disagreement. Whilst a segment of the lifeworld can become problematic, and acquire the status of contingent reality, it is not possible that everything could be called into question at once.

In analysing particular social phenomena from a lifeworld perspective, we adopt the perspective of participants, acting subjects recognising and responding to norms (TCA2, 132). From this perspective, the lifeworld itself remains withdrawn from thematisation and unavailable to theory. In order to explain the reproduction of the lifeworld as a whole, Habermas turns to the lifeworld’s structural components, the functions that the medium of language fulfils in the lifeworld’s reproduction. This functional account of structural components distances us from the particularity of a participant’s point of view. There are three structural components. Firstly, in enabling mutual understanding, language functions to transmit and renew cultural knowledge, allowing reproduction of the contents of culture. Secondly, language functions to coordinate action, thus integrating society and stabilizing group solidarity. Thirdly, language functions to form personal identities by the socialisation of responsible social actors. Language thus serves to connect new situations with present conditions in the lifeworld dimensions of culture, society and personality (TCA2, 137-8).

The lifeworld concept thus refers to a network of cultural assumptions enabling a common understanding of situations, institutions and societal norms underpinning social solidarity, personal self-understandings, competencies and skills. The three structural components of the lifeworld are reproduced through the symbolic interactions of everyday communicative practice in cultural reproduction, social integration and socialisation. This division maps onto the three formal world concepts and their associated validity claims. Thus “culture” is a resource to challenge and redeem claims of truth in the objective world, whilst “society” is a resource to redeem normative rightness within the inter-subjective world and “personality” a resource to challenge and redeem claims of truthfulness and authenticity of the subjective world. All three structural components reproduce society by operating in a circular fashion, mediating between the particularity of the individual or situation and the general framework of lifeworld resources (TCA2, 141-4). The failure of culture, society and personality to adequately reproduce themselves results respectively in, the loss of meaning and rationality, anomie (loss of normative motivation) and mental illness (TCA2, 140-1).

Into this account, Habermas introduces a social evolutionary dimension which I will elaborate shortly. For now, the lifeworld should not be seen as a conservative “given” but developing towards
increasingly rationality. Emerging out of unquestioned social solidarity, societies increasingly
differentiate claims into relevant validity spheres of truth, rightness and truthfulness pertaining to
objective, inter-subjective and subjective “worlds”. This leads to the increasing distinction between
the structural components of the lifeworld, increasing abstraction and separation of form and content
(TCA2, 146). The increasingly important role played by science in the lifeworld reflects these
developments.

**Systems**

There is a cost to this increasingly differentiated rationality. As actions become less subject to
guidance by compulsory norms, both the volume of communication required and the risk of failure to
reach consensus increase. As social integration becomes more difficult, systems such as the market
economy develop, relieving the burden on communicative action (TCA2, 150). Whilst the lifeworld
reproduces society communicatively, systems specifically reproduce the material aspects through the
purposive activity of work, by which individuals realise their aims.

Habermas regards the systems perspective as a necessary complement to the lifeworld perspective
which on its own is blind to materially induced distortions of symbolic reproduction (TCA2 148-50).
Including a systems perspective alongside, but distinct from, a lifeworld perspective, enables two
analytic perspectives necessary for social theory (Honneth & Joas, 253). Social integration relies on
lifeworld mechanisms that harmonise action orientations by normatively secured consensus. System
integration relies on the various sorts of “invisible hand” mechanisms through which the pursuit of
individual interests produces desired patterns, unintended and often unrecognised by actors (TCA2,
150). For example, the market economy stabilises non-intended consequences of actions by way of
functionally intermeshing action consequences, ideally matching availability of resources to levels of
needs. Systemic outputs (prices) are also inputs which push systems towards a homeostatic
equilibrium, a “goal state”. Attempting to avoid socially unacceptable outcomes that market processes
induce, governments intervene in this feedback loop, by a “steering medium”, such as power or
money. This encodes “information” into systems and “a purposive rational attitude toward calculable
amounts of value, so enabling a generalised, strategic influence on the decisions of other participants
while *bypassing* consensus-oriented communication” (TCA2, 183). Scientific discourses, for example
economics, clearly have a pre-eminent role in constructing, maintaining, adjusting and analysing
systems, as well as, with their translation and dissemination into the lifeworld, the anchoring role of
legitimation.

Habermas maintains that by transforming the systems and lifeworld perspectives into each other, we
see that the more complex systems become, the more provincial lifeworlds become. The more the
lifeworld is expanded and differentiated, the more its analysis needs to be complemented by a systems
analysis to get at the counterintuitive aspects of sociation. Goal-directed activities of social actors are coordinated not only by their intended orientations to mutual agreement but also by functional interconnections not intended or recognised within the horizon of everyday practice. The problem for social theory is how to connect the two conceptual strategies.

Habermas insists that symbolic reproduction of the lifeworld remains a matter for communicative action.\textsuperscript{114} Whilst non-linguistic mechanisms can free people from the need for excessive negotiation and the risk of consensus failure, this freedom can't be extended to areas of life that primarily fulfil functions of cultural reproduction, social integration and socialisation through the process of mutual understanding in language. The structure of mutual understanding is required to recognise the criticisable validity claims embedded in lifeworld contexts defined by cultural traditions, institutional orders and competences. Communicative action feeds off the resources of the lifeworld by actualising rationality potential as particular shared lifeworld contexts lose their background quasi-naturalness by becoming subject to validity claims.

Systems theory doesn't make the distinctions necessary to grasp the pathologies that emerge in modernity as pathologies. Increased complexity achieved at the cost of the lifeworld does not appear as a cost. From the systems perspective, society appears as a self-regulating system, analogous to an organism. But the systems analyst can only judge whether disequilibria reach a critical point if survival limits can be identified. Unlike organisms, there is no obvious sense of limits to social systems. The social scientist can speak of crises only when relevant social groups themselves experience systematically induced structural changes as critical to their continued existence and feel their identities threatened. In the case of social, rather than natural evolution, the criteria of what constitutes a good society are open to the ongoing determination by social members. Societies need to be grasped hermeneutically, from the internal perspective of participants.

However, a social theory that sees society totally absorbed into the lifeworld is also problematic. It offers only an internal perspective, where the lifeworld that members construct from common cultural traditions is coextensive with society. This perspective regards social actors as completely autonomous, able to rationally orient themselves to criticisable validity claims. Yet the web of interactions extending across social groups and throughout history cannot be explained by the conscious intentions of actors who can't fully control either the possibilities for mutual understanding and conflict, or the consequences and side effects of their actions. From the lifeworld perspective, the

\textsuperscript{114} Habermas is critical of Parsons' attempt to reduce the lifeworld to a subsystem, by introducing the notion of inter-systemic relations regulated by media modelled on the medium of money. Parsons is over-generalising the notion of media since only material reproduction can take place via non-linguistic media (TCA2, 265-82).
counter-intuitive aspects of systematically distorted societal reproduction are not in view (TCA2, 150-1).

Habermas’ social theory, combining both lifeworld and systems analyses, aims to account for the relations, interchanges and influences between these two levels. However, only the lifeworld can define the social system as a whole. Habermas insists on the priority of the lifeworld since the very objects of systems theoretical concepts presuppose their prior identification from a lifeworld perspective (TCA2, 151). Ultimately, systems remain anchored to the lifeworld on which they depend for legitimation, despite the dependence of lifeworlds on increasingly autonomous systems to relief pressures and risks of communication. Habermas describes two sorts of "relief mechanisms" that arise as systems emerge from the lifeworld. “Replacement” of communication involves media such as money or power which replace linguistic understanding “empirically” by motivating sanctions, rewards or threats (TCA2, 183). Such delinguistified media connect up interactions into increasingly complex networks that no one can of fully understand, or for which is required to accept responsibility. Whilst serving to increase the scope of freedom of action oriented towards success, this mechanism can also result in massive disfunctionality, as evidenced by the Global Financial Crisis. Money and power bypass processes of communication oriented towards agreement by signals which change agents’ incentives and behaviours (TCA2, 183).

“Condensation” of communication involves media such as “professional reputation” of a scientist which is “in the last instance” rationally accountable (TCA2, 185). Professional reputation achieves consensus by specialising in certain aspects of validity and hierarchising processes of agreement. People give more weight to the opinions of scientists within their area of expertise and generally accept non-contentious scientific claims without requiring full justification. This sort of relief mechanism is rational. Whilst persons having professional reputation certainly have “private” interests, these interests ideally operate as neither inducements nor sanctions to produce particular opinions. Rather, distorting private interests tend to be screened out by the institutional norms and conventions of science. Instead of uncoupling interaction from the lifeworld context, condensation remains a second-order processes of consensus formation in language. I shall return to Habermas’ notion of condensation, to incorporate it into a discussion of the closely related, but more threatening, idea of “splitting-off” of expert knowledge.

3. **EVOLUTION and RATIONALISATION of SOCIETY**

Before considering Habermas’ diagnosis of modernity, it is necessary to discuss his theory of social evolution as a basis for his normative standpoint. This theory traces the acquisition of communicative
competences, as learning capacities which respond to internal and external challenges.

Communicative competence is not given, but acquired by an identifiable ontogenesis, which Habermas extends to a developmental history of societies. Much of this background is assumed in TCA, so we will need to backtrack to Habermas’ earlier Communication and the Evolution of Society (hereafter CES). Borrowing from Piaget’s structuralist account of ego development and Kohlberg’s work on moral competence, Habermas hypothesises a number of developmental stages of societies, premised on the analogy with children, who proceed through an irreversible sequence of learning stages, with later stages incorporating and modifying elements of earlier ones. He thus reconstructs the developmental logic of the learning processes of societies by relating forms of societies to developmental sequences adapted from Piaget’s and Kohlberg’s models of individual development (CES, 156-8).

This staged development involves a process of “decentering”, a term borrowed from Piaget, suggesting stepping back from the immediacy of a situation to adopt a broader perspective (TCA1, 69). What is taken for granted at one level becomes the object of critical reflection at the next. Having acquired the basic distinction between the lifeworld and the world, the mature individual is able to question the taken-for-granted assumptions of the lifeworld (MCCA, 138). This requires the ability to differentiate between cognitive, normative and expressive speech acts and treat one’s assumptions as hypothetical by suspending the binding force they exert on one’s beliefs. Developmental stages do not map onto social processes in any mechanical way, but are speculative accounts of social progress, open to empirical investigation. Habermas remains cautious of too easy an assimilation between individual and social development (CES, 98). Whilst the patterns of individual developmental psychology may turn up in social evolution, this offers only a tentative guide to a social progress that moves from particularistic to universal (CES, 99). Habermas’ claim is that societies “learn” in a derivative sense. Only a framework for learning is provided – a “logic” within which individuals learn to deal with conflicts and solve problems (CES, 121). The “dynamics” of the lifeworld’s actual development are influenced by forces arising from problems in society’s material reproduction, whilst constrained within the developmental logic of structural possibilities (TCA2, 148). By this logic, Habermas establishes a normative standpoint, from which deviations are visible.

Applying this social evolutionary theory, Habermas sees the lifeworlds of archaic societies as having a high degree of congruence in relations between social institutions, cultural worldviews and persons (TCA2, 156). In “the magical-animalistic representational world of paleolithic societies … all natural and social phenomena were interwoven and could be transformed into one another” (CES, 104). Categorical distinctions between society and natural surroundings tend to be blurred. These homogeneous lifeworlds become rationalised by the process of “linguistification of the sacred”, the ongoing development of communicative competence which differentiates the lifeworld’s structural
components (TCA2, 77). This process represents successive releases of communicative potential by a de-centered learning process that allows a critical distance to open interpretations to discursive validation.

Habermas’ account of the progressive unfettering of rationality enables a developmental ordering according to the degree to which aspects of validity have been differentiated from one another. At one end, stands ritual practice, at the other, the practice of argumentation. Throughout the course of social evolution, action oriented to mutual understanding gains increasing independence from normative contexts provided by traditional, religious or metaphysical worldviews. Institutionally required value orientations become general and formal, law and morality universal, and individuals autonomous. Consensus formation takes place according to the strength of argument governed by abstract rules. Language takes over the functions of achieving understanding, coordinating action, and socialising individuals, thus becoming the medium of cultural reproduction, social integration and socialisation. As lifeworlds become rationalised, systems arise to relieve pressures on the linguistic resources of societies. Systems are initially uncoupled from the lifeworld by differentiation between coordination through communicative action directed at intentions and coordination through functional interconnections of action consequences. Society, in the course of social evolution, becomes differentiated both as a system and as a lifeworld. Systems increase in capacity to autonomously regulate material reproduction behind participant’s backs, whilst lifeworlds increase in rationality by the separation of culture, society and personality. This progressively rationalised lifeworld becomes both uncoupled from, and dependent upon, increasingly complex, formally organised systems of action, like the economy and state administration (TCA, 153-4). What Habermas originally treated as a methodological distinction, he now regards as an ontological distinction between lifeworld and system, uncoupled in modern capitalist societies (Honneth & Joas, 255).

Whilst Habemas generally views the development of systems positively, as a necessary complement to rational progress in the lifeworld, he also recognises their complexity and opacity as a threat. As the market economy emerged, money became an inter-systemic medium, serving to regulate exchanges with the social environment and weakening the primacy of the lifeworld (TCA2, 171). Success-oriented action had previously been bound to norms of action embedded in communicative action within the cooperative social task-oriented framework of traditional societies. Now with money as a medium, success-oriented action can be steered by egocentric calculations of utility, calling for “an objectivating attitude even in regard to interpersonal relations” (TCA2, 196). Societies are steered by mechanisms such as money and power in a manner largely disconnected from norms and values of the lifeworld. Subsystems of purposive-rational economic and administrative action become independent of their moral-political foundations. However, the tendency towards ever-expanding bureaucratic and economic subsystems is by no means unavoidable. Whilst processes of mutual understanding are
partly suspended in formally organised domains, the steering media of power and money remain anchored in the lifeworld (TCA2, 311-2). With this we come to Habermas’ diagnosis of modernity.

4 THE DIAGNOSIS OF MODERNITY

In Habermas’ acceptance speech at the Johann Wolfgang Goethe University, he pleaded for a continuation of a modernity enlightened by reflection on its aporias and failures. He characterised modernity as an ongoing historical phase in which complex societies emerge, characterised by diverse conceptions of life. Modernity’s rationalised lifeworld is increasingly differentiated, leading to demands for reasons, the intellectualisation of life and the calculability of things. As well as the growth of systematic mechanisms such as the economy and bureaucracies, modernity has yielded modern science, positive law, secular, principle-governed ethics and democratic forms of will-formation. Whilst Habermas thinks “that something is deeply amiss in the rational society in which [he] grew up and now live[s]”, TCA attempts to defend these achievements (Habermas 1986, 126).115

Habermas’ diagnosis of modernity is presented as an analysis of the insights and blind spots of previous theorists – Marx, Durkheim, Mead, Parsons, Weber, Horkheimer and Adorno. Whilst Habermas credits Marx with recognising the distinction between lifeworld and system, Marx failed to distinguish between rationalisation of the lifeworld, whereby the co-ordination of a complex modern society necessarily passes to systemic steering mechanisms, and colonisation of the lifeworld, whereby non-symbolic steering media invade the lifeworld, reifying what were communicatively structured domains. He thus failed to appreciate “the intrinsic evolutionary value that media-steered subsystems possess”, such as the rationalisation of economic and legal systems which open new possibilities of freedom by enabling increased communicative rationality, whilst liberating actors from its excessive demands (TCA2, 339). Weber makes the mistake of explaining the ills of modernity as the inevitable outcome of disenchantment, which he equates to the growth of purposive rationality, thus failing to discern that critical reflection enables distanciation from the parochial nature of particular views lodged in the lifeworld (TCA1, 72, 236). The threat to the lifeworld is not rationalisation per se, but rationalisation enabling “the emergence and growth of subsystems whose independent imperatives turn back destructively upon the lifeworld itself” (TCA, 186). The effects of this colonisation appear “at the seam between the system and the lifeworld”, along which new potentials for emancipation and resistance develop (TCA1, xxxv).

115 Habermas’ commitment to the progressive nature of modernity’s “unfinished project” contrasts with Foucault’s methodological agnosticism towards progress, leading Habermas to characterise him, unfairly, as an anti-modernist “young conservative” (Habermas 1996, 53).
Habermas’ critique is not directed at capitalism per se, but the monetisation of the lifeworld which involves a form of system rationality that has become autonomous. If money only circulates within the economic system, its function as mediator between supply and demand of material goods is not necessarily problematic. However, when the economic system penetrates into communicative practices of everyday life, it reifies social relations. Reification also derives from the political system which seeks mass loyalty to secure its own continuation. Comprehensive political strategies aim to prevent and manage crises by raising practical-political questions to the status of special problems to be dealt with by experts, so short-circuiting genuine public debate (TCA2, 350). Habermas wants to work out these tensions between capitalism and democracy in a way that establishes the primacy of democratically legitimised norms.

Habermas’ diagnosis focuses on two inter-related tendencies connected with disturbances in lifeworld reproduction – “cultural impoverishment” and the “colonisation of the lifeworld”. Cultural impoverishment is caused by the splitting-off of elite subcultures, resulting in “fragmented consciousness” incapable of integrating cognitive, normative and aesthetic understandings of reality in a critical way (TCA2, 355). Noting the tendency of specialised technical expertise to lead to centralisation, hierarchical management, conformity and routinisation, Habermas shares Foucault’s concern about the “normalisation” induced by discipline and biopower, which robs citizens of their critical capacities. The colonisation of the lifeworld involves substituting strategic forms of economic and legal action mediated by money and power for communicative forms of action responsible for socialisation, cultural transmission, and social integration. Both tendencies reflect the imperatives of capitalist modernisation which restructures everyday life in accordance with systems requiring a hierarchical specialisation of labour, reducing subjects to objects of mechanical processes of production and consumption.

According to Habermas’ account of rationalisation, we would expect the lifeworld to be capable of resisting systemic imperatives by demanding reasons which call systems to account. In modernity, as the cultural value spheres become increasingly differentiated, disenchanted culture should lose any capacity for ideology (TCA2, 352). Culture should be too open to demands for reasons, “too transparent to provide a niche for [the] structural violence [of systems imperatives] by means of inconspicuous restrictions on communication” (TCA2, 196). Competition between systems integration and social integration should become more visible (TCA2, 355). But this doesn’t happen because holistic interpretations “never attain that level of articulation at which knowledge can be accepted as valid by the standards of cultural modernity. Everyday consciousness … becomes fragmented.” The lay person loses authority in areas of knowledge which have become the exclusive domain of expert specialists. In effect, capitalism has found “a functional equivalent for ideology” by
alienating us from our critical aptitudes (TCA2, 355). To be critical we must be able to synthesise different validity aspects. Modernity forces us to bracket areas of knowledge accessible only to specialists (TCA2, 355). Thus deformations of the lifeworld “ought not be attributed either to the rationalisation of the lifeworld as such or to increasing system complexity as such …[but to] the elitist splitting-off of expert cultures from communicative action in daily life”, leading to impoverishment of communication and dying out of vital traditions (TCA2, 330). The resulting fragmented consciousness lacks the holistic resources to recognise and articulate resistance to systemic imperatives, the effects of which are nonetheless strongly felt. Pathologies develop across the structural components of the lifeworld - loss of meaning, withdrawal of legitimation, crises in orientation and education (cultural reproduction), insecurity in collective identity, anomic, alienation (social integration), breakdown in tradition, withdrawal of motivation and psychopathologies (socialisation) (Thompson & Held, 280).

Today science, morality and art are autonomous spheres dealt with by specialists. Unable to reach consensus “across the whole spectrum”, we experience a loss because in everyday life, “cognitive interpretations, moral expectations, expressions and valuations have to interpenetrate and form a rational interconnectedness via the transfer of validity that is possible in the performative attitude” (TCA2, 327). To resolve real life situations in which different value spheres are inextricably compounded, we require the competence to analyse, discuss and balance different validity claims, rather than handing them over to expert specialists with partial perspectives. But with the splitting-off of expert knowledge, the lifeworld’s competencies are weakened, enabling its colonisation by systems which substitute strategic forms of economic and legal action, mediated by money and power, for communicative forms of action responsible for socialisation, cultural transmission, and social integration. I will now address both tendencies identified by Habermas’ diagnosis and, in relating them to science, point to some difficulties.

5 FRAGMENTED CONSCIOUSNESS

The elitist splitting-off of expert cultures is clearly true of modern science. Since the 19th century, the sciences have excluded the amateur and developed specialised vocabularies that exclude communicative dimensions extraneous to their purposes. This enables inter-subjective agreement within the scientific community and precision and economy within scientific practice. It enables the exclusive focus on truth within restricted object domains. Scientific disciplines enforce boundaries between themselves and other scientific and non-scientific disciplines. Whilst the lifeworld seeks to draw everything that can be grasped by reason into a comprehensive network, science pulls back to become increasingly inaccessible.
Habermas firstly addresses this inaccessibility in terms of the relief mechanism of “condensation” which I have already mentioned. Condensation is particularly relevant to the dissemination of scientific discourses into a technologically expanded public realm which has freed speech from its parochial spatial and temporal limits. According to Habermas, whilst claims based on scientific authority or reputation are widely dispersed they are still, in principle, rationally redeemable. Although de-contextualised by transmission far beyond their original contexts, these claims ultimately constitute undertakings to provide reasons. Such scientific claims “relieve interaction from yes/no positions of criticisable validity claims only in the first instance ... [but] in the last instance, remain dependent on the actions of responsible actors” (TCA2, 185).

At this point, Habermas doesn’t acknowledge the limits to rational accountability within a mass society which no longer relies on face-to-face contacts. By claiming that within specialised scientific discourses, individuals are responsible for providing reasons “in the last instance”, Habermas gives insufficient weight to the distorted nature of communication in an attenuated public realm, and the contingency of the contexts which frame communicative uses and meanings. Whilst scientific reputation is a resource that in principle enables redemption of claims, in actual fact many claims are not adequately redeemed. Scientific facts frequently appear in a public realm completely isolated from the scientific research practices and programs which provide the only contexts against which their validity can be determined and their limits recognised. Habermas was aware of this in his earlier work, referring to “the fifty thousand scientific journals that are published today” (1968) and commenting that “it is increasingly difficult to survey, order and process these in order to get an overview” (TRS, 77). Today, fifty years later, we see significant political effects of this increasing difficulty in the hopelessly inadequate contextualisation of isolated scientific facts in discussion about climate change, as it enters the lifeworld of non-scientists.

In contrast to Habermas’ idealised picture of condensation as rational justification “in the last instance”, the notion of “fragmentation of consciousness” marks a threshold beyond which knowledge has been irredeemably “split-off”. This notion modifies Weber’s thesis of the separation of value spheres. These value spheres, differentiated by the cultural rationalisation of the lifeworld, represented for Weber irreconcilable and competing spheres of rationality. Habermas wants to endorse this differentiation as a progressive cultural development. The substance of particular value spheres is still open to judgment according to universal criteria, by virtue of the unity of form of each type of validity claim. This is what Habermas sought to establish in formal pragmatics (TCA1, 249). Accordingly, the ills that Weber identified should not be attributed to conflicting value spheres. Habermas wants to drive a wedge into Weber’s account of disenchantment by insisting that “differentiation” itself is not the problem, but “splitting away” from the lifeworld is. If only differentiated, specialised spheres can
still be “ultimately” brought together within the lifeworld’s “unity of reason”, where three irreducible dimensions of validity can be raised and redeemed by shared communicative norms. But by becoming the exclusive domain of expert specialists, specialised scientific discourses are “split-off”, unable to be brought within the ambit of the lifeworld’s communicative resources.

Habermas’ account of fragmentation remains undeveloped. He doesn’t consider the nature of “trade-offs” between irreconcilable claims which reflect only strategic consensuses. This can be seen by considering the role of science in, for example, an Environmental Impact Statement that addresses impacts of uranium mining on the natural and social environments, the economy and indigenous heritage. In this case, there are no rationally compelling criteria, beyond those of legal proceduralism, on how to adjudicate between the irreconcilable values which are unavoidable consequences of disenchanted modernity. In such cases, the authority of scientific-technical discourses, “split-off” from the lifeworld, is drawn upon to enable practical decisions. Such discourses tend to bracket subjective and qualitative aspects that defy quantification. In so doing, scientific experts adopt an observer’s perspective towards matters imbued with meanings, significances and values only available from the participant’s first-person perspective. These meanings, significances and values can become objects of investigation by further scientific discourses, directed at gaining knowledge from a more universal standpoint. By this distanciation, what is meaningful, value-laden and significant, is increasingly reduced to an abstract “resource” to be weighed against other resources. Whilst this simplification enables decisions according to democratic legal norms, it is at the cost of diminishing what is meaningful, value-laden and significant.

We moderns both endorse and lament this situation, since we have no alternative but to demand that, for example, economic values must somehow be “balanced” against aesthetic, traditional or cultural values, in a consistent, accountable and transparent manner. Faced with disconcerting irreconcilability, and wanting to avoid conflict, we turn to our most authoritative knowledge to provide an answer beyond reproach. But for science to grant its authoritative support to a technological proposal with far-reaching consequences, its investigations and findings must fit within legal categories generated by the competing imperatives of capitalism and democracy. In this role, scientific discourses are granted a practical authority that science itself doesn’t possess. Whilst this doesn’t reduce such decisions to irrationality, it does suggest a rather thin formal procedural-legal rationality. I am not denying that science is essential to understanding and assessing technological proposals and assisting decision-making in a way that avoids open conflicts. What I want to highlight is that scientific discourses, by virtue of the authority granted them, appear to justify political decisions, and this appearance stifles deliberation. Whilst the fact of science splitting-off to become an elitist subculture is certainly important, it is not clear how this fact bears on what seems better described as this ideological role.
Although Habermas’ shift to the intersubjective communicative paradigm undoubtedly represents a theoretical gain, I would suggest that the ideological role of his earlier “technological consciousness” better captures the sense of authority granted to science than his discussion of the “impoverishment of culture”. The elitist splitting-off of science is not a problem merely by virtue of inaccessibility. What the notion of inaccessibility doesn’t capture, and is better captured by Habermas’ earlier critiques of positivism, is the precise nature of a misevaluation of science which sees it justifying practical decisions as matters of objective exigency. This mis-evaluation cannot simply be passed off as either “splitting-off of expert knowledge” or manipulation by powerful interests, although both are part of the story. What is more significant to the relation between science and society is the role, and consequences, of this ideological interpretation of science in persuading the polity to accept political decisions. Whilst interpretations of science as able to provide authoritative justification, serve the social function of consolidating consensus around practical decisions taken in accordance with legal democratic norms, such consensus is achieved only at the cost of narrowing discussion and deliberation.

I now want to go beyond Habermas’ analysis to argue that by their very nature, scientific discourses lend themselves to uses which enable this persuasion. This is seen not only by the role of scientific discourses within the formal arena of formulating and interpreting law, but also within the informal arena of will-formation within the public sphere.\(^{116}\) We may be tempted to insist that scientific discourses are not themselves ideological. Alternatively, we may entertain Foucault’s view of ideology and science as inextricably bound together.\(^{117}\) Regardless of such nuances, I want to claim that scientific practice has inherent features which lend science to ideological use in authoritative affirming or undermining the legitimacy of practical decisions. One such feature can be seen by the relation between established scientific theories and facts which seem to contradict these theories.

The scientific community, confronted with anomalous facts, does not simply reject theories which otherwise have explanatory and predictive power and are validated by separate lines of evidence. Theories are only rejected when these facts stubbornly persist and continue to accumulate and better theories seem possible. Kuhn described this “paradigm shift” as akin to a religious conversion, not meaning that science is irrational, but rather that there are no criteria by which the scientific community is compelled to shift at any particular point (Kuhn 1996). Rather than logical compulsion, it is a matter of judgement which involves argument within the scientific community. The criteria by

\(^{116}\) I will elaborate on these two arenas in my discussion of BFN in Chapter 7.

\(^{117}\) See my discussion of Foucault’s notion of the “tactical polyvalence” of discourses, by which the same discourse can be used for different tactical purposes. (Ch. 3)
which scientists accept something as “true” is partly aesthetic (scientists talk about the “elegance” of solutions), simplicity (Ockham’s razor) and pragmatism. Climate science, which is primarily statistical, seems a case where pragmatism has purchase, where scientists might reasonably argue “certainly there are a few facts that we can’t yet explain, but given what the theory does explain and what is at stake, we should treat it as true.” In other words, there is a certain unspecifiable point at which we need to settle the irritation of doubt and fix our beliefs, although this point can’t be specified.

This indeterminacy is played out in the adversarial nature of scientific publication which results in the public dissemination of facts stubbornly resistant to incorporation into theory. This is unproblematic within the language games of the scientific community in which such facts are viewed within the context of scientific discourse. However, when scientific discourses migrate into the lifeworld of non-scientists, these anomalous facts are frequently understood as sufficient by themselves to undermine theories accepted by the scientific community. The social use of such decontextualised facts by power borrows the authority of science which, at the same time, is reinforced. Habermas is aware of this. “The political consequences of the authority enjoyed by the scientific system are ambivalent … short-lived popular syntheses of isolated pieces of information, which have taken the place of global interpretations, secure the authority of science in abstracto” (LC, 84) This authority assumes an ideological role when it licences facts, isolated from any context by which they can be adequately understood, or their significance judged, to bolster support for particular power interests which claim to coincide with undisputable positive outcomes. Habermas doesn’t develop this thought.

To elucidate the ideological role of science, I have gone beyond Habermas’ arguments on “condensation” and “fragmented consciousness” to suggest that scientific knowledge lends itself to distortions of communication and understanding. Seeking to steady our uncertainty, we turn to the authority of science to provide answers which it cannot really provide. Since what is at stake demands a decision based on unimpeachable reasons, we are vulnerable to the appeal of one-sided perspectives that either give excessive authority to science, or seek authority elsewhere. The problem is not only “splitting-away”, but the misevaluation of science as the authoritative ground that can assuage the demands for certainty, reassurance and reconciliation we seek for practical questions. The fact that these demands are located within the lifeworld, vastly complicates Habermas’ distinction between lifeworld and system, whereby the lifeworld is the home of communicative resources which are distorted by systematic imperatives from the outside.
Habermas’ powerful metaphor of colonisation suggests that the lifeworld, weakened by fragmentation and loss of internal cohesion, is no match for systems which, barely comprehended, invade to restructure its resources according to their autonomous imperatives. “When stripped of their ideological veils, imperatives of autonomous subsystems make their way into the lifeworld from the outside – like colonial masters coming into a tribal society – and force a process of assimilation on it. The diffused perspectives of the local culture cannot be sufficiently coordinated to permit the play of the metropolis and the world market to be grasped from the periphery” (TCA2, 355). The sciences, through which systems are constructed, comprehended and legitimated, have been subsumed within areas of expertise which exclude the comprehensive perspectives of the lifeworld which, now fragmented, is restructured by systemic imperatives.

Whilst accepting the fact of reification in advanced capitalism, Habermas doesn’t accept Horkheimer’s and Adorno’s totalising critique of instrumental reason which leaves no position from which reason can speak (Adorno & Horkheimer 1997). Whilst colonisation marks a threshold beyond which systems, originally anchored to the lifeworld but now adrift, return to damage the lifeworld, this process only reflects the dynamics of development - that is, contingent empirical factors. It does not reflect the logic of development - the learning process. Rationalisation could have been balanced, not one-sided as it has turned out under advanced capitalism. It could have, and still might, allow the democratic normative control of subsystems and their steering mechanisms (Habermas 1990, 109).

Habermas’ colonisation thesis must be examined within the context of the developments of contemporary advanced capitalism, characterised by the welfare state. 118 According to his account, the welfare state largely pacified class conflicts arising from acute economic crises, by manipulating conditions around the margins to avoid social impacts whilst not restricting the market (TCA2 350-1). Habermas accounts for government interventionism, the welfare state and mass democracy by analysing multiple interchange relations between lifeworld and the economic and administrative subsystems. The economic subsystem exchanges wages for labour, and goods and services for consumer demands. The administrative subsystem exchanges organisational achievements for taxes and political decisions for the mass loyalty of voters. These exchange relations define new social roles within the lifeworld (TCA2, 343). Freedom is lost by the dependency of these roles on their respective subsystems and the limitations that the identification with these roles entails. Meaning is lost by the

118 Although the welfare state has been under challenge since the 1980’s, it is far from disbanded, and remains as a paradigm, though one only partially realised. For this reason Habermas analysis of colonisation in terms of the welfare state, though requiring some updating and adjustment, is hardly obsolete.
tendency toward abstractions, which occur “whenever the lifeworld, in its interchanges with the economic or administrative systems, has to adapt itself to steering media” (TCA2, 322). In the administration of the welfare state, all clients are treated by the same formal procedures. This one-sided rationalisation shows in everyday practice in which consumerism, achievement motivation and competition gain increasing dominance over moral-practical orientations.

Habermas discusses colonisation in terms of “juridification” - the increasing tendency towards formal law in modern societies. Juridification opens up informally organised areas to bureaucratic interference and judicial control. Juridification is ambiguous. In its latest stage, protections – collective bargaining, wage awards, social insurance - are experienced as both liberating, enabling citizens to exercise rights, but also as alienating. The legal processing of the concrete situations of individuals’ lives must be “administratively digested” by being forced into abstract terms. What Habermas is talking about here relates to Foucault’s notion of “fabricating” people discussed in Chapter 3. Juridification relies on systems of classifications, social ontologies developed and given authority by human, social and medical sciences. Exactly what it is to be a refugee, a dependent, a delinquent, is determined by legal definitions which borrow from these sciences to define and explain, so lending the classification both objectivity and authority. The criteria, for example, by which someone qualifies for a certain relation with the administration can often be traced back to scientific discourses which authorise and justify administrative actions and decisions. It is not only the fact that scientific categories serve instrumental purposes of classification but, as Foucault shows, they feed into socialisation processes in a way that socially constructs human types, crafting dependencies and reconfiguring lifeworlds behind our backs.

Habermas’ colonisation thesis also accounts for recent conflicts which deviate from the usual institutionalised conflicts over distribution. These new conflicts are not about domains of material reproduction and are no longer channelled through political parties. They arise in domains of cultural reproduction, social integration and socialisation concerned with the restoration of damaged ways of life. The "new politics" of alternative movements turns on quality of life, self-realisation, participation and human rights. It includes environmental, peace, and squatter movements, struggles for regional, linguistic and religious independence. This phenomenon demonstrates the aptness of the colonisation metaphor, suggesting subjects within a threatened lifeworld can resist, reforge identities or gain independence.119

Habermas’ colonisation thesis has been criticised for implying a clean divide between colonised and coloniser and a one-way movement of systems colonising the lifeworld (McCarthy 1991, Ch.6; Fraser

119 What Habermas identified as colonisation, Foucault identifies as normalisation in terms of biopower, which is similarly resisted and subverted by counter-movements. See Ch 4
For example, in relation to the coloniser/colonised division, Habermas sees organisational communication falling under the premises of formally regulated domains of action, such that "communicative action forfeits its validity basis in the interior of organisations" (TCA2 310). Habermas overstates the situation. Organisations constantly establish, review, renew, reinterpret and revise agreements. He is not correct to claim that by "challeng[ing] the assimilative powers of an all-encompassing lifeworld" the economy and bureaucracy “congeal into a second nature of a norm-free sociality that can appear as an objectified context of life” (TCA2, 173). It is not clear that systems theoretic concepts are useful here. Whilst it may be plausible to say that the market, as an ethically neutralised system of action, is “norm-free”, bureaucracies can’t be totally norm-free. If organisations are systems, they are porous to the lifeworld.

By confining his discussion of power to systems contexts, and thus implying the one-way movement of colonisation, Habermas appears to screen power out from the lifeworld. Whilst he doesn’t think of the lifeworld as free from power, his analysis doesn’t address the significance of socialisation which involves forms of domination reproduced in the modern lifeworld. Fraser argues that the family perpetuates relations of oppression as much as it reproduces values and cultural norms. To be truly critical, critical theory must be capable of foregrounding this domination (Fraser 1989, 138). By making firm institutional divisions, such as between families and public sphere on one side and business and state bureaucratic organisations on the other, Habermas tends to portray systems as the locus of power which colonises the lifeworld. Yet families are responsible for material reproduction of society just as business and state organisations are responsible for symbolic reproduction. And influence appears to go both directions.

In response to such criticisms, Habermas softened the lifeworld/system distinction, emphasising that all social phenomena can be viewed from both lifeworld and system perspectives. The economy and the state, he now tells us, are “primarily” integrated systemically. Rather than the distinction between lifeworld and system being based simply on the type of action (instrumental or communicative), Habermas now sees the difference in the dominance of one action type and media within a certain domain (Honneth & Joas, 254). The distinction is not of the nature of an “either/or” but rather a “more or less”. Habermas grants that the lifeworld “by no means offers an innocent image of power-free spheres of communication” and concedes the possibility of a reverse movement which subjects systems to the normative restrictions of the lifeworld (Honneth & Joas, 254).

120 For example, in a transaction, the medium of money coordinates systemic effects behind my back. But in the very same action, I have undertaken a communicative interaction by which both parties are oriented towards reaching mutual understanding.

121 A reverse movement could subject systems to the normative restrictions of the lifeworld by, for example, the boycotting of products for reasons directed towards remote functional consequences.
which side imposes limitations has to be treated as an empirical question which cannot beforehand be
decided on the analytical level in favour of systems ... [the] colonisation of the lifeworld and the
democratic control of the dynamics of systems ... represent two equally justified analytical
perspectives” (Habermas 1990, 109). It may well be that systems currently colonise the lifeworld, but
there is nothing to suggest that this is necessary.

In terms of Habermas’ thesis, scientific discourses could be considered to operate as links that
formulate, constitute, describe, interrogate and justify systemic power to the lifeworld. On one hand,
the condensation thesis suggests science remains “ultimately” accountable, and systems ultimately
anchored, to the lifeworld. On the other hand, the fragmentation thesis suggests a threshold has been
crossed, beyond which science is irredeemably split-off from the lifeworld, such that systems cannot
be held to account by the lifeworld. Whilst this ambivalence suggests that Habermas is alerting us to a
tendency, rather than an already accomplished situation, this account needs to be expanded. He could
examine the role of science in the juridification processes through which colonisation operates, since
these processes explicitly adopt and apply elements of scientific discourses. The social function of
scientific discourses within juridification doesn’t automatically arise from their truth, but from their
being granted authoritative role in the face of challenges to the legitimacy of juridical and
administrative procedures. Since the sciences are accepted as the best form of knowledge available,
this authority serves to maintain social cohesion, avoid conflict and generate consensus. It settles the
irritation of doubt, closing off questions to endless deliberation, providing the certainty of a reassuring
framework which we can collectively accommodate. Systems’ autonomy is authorised by science
which addresses the lifeworld’s need for certainty and reassurance. Such psychological-functional
accounts can unravel some of the complex threads between system and lifeworld that Habermas
neglects.

I now want to return to Habermas’ earlier view of scientific technology (discussed in Chapter 5)
according to which the ills of modernity result from an imbalance, whereby instrumental objectifying
reason comes to dominate “interaction”, seemingly implying the neutrality of scientific technology “in
itself”. As we have seen, Habermas replaced KHI’s schema of anthropological interests with his
communicative paradigm to yield universal categories of the objective, intersubjective and subjective
“worlds” with their corresponding validity claims, attitudes and cultural spheres. The problem is that
any schemata risks separating what is only analytically separable but actually compounded.
Following his debate with Marcuse, Habermas retained the view that “there is for this domain of
reality [the objective world] only one theoretically fruitful, attitude, namely the objectivating attitude
of the natural-scientific experimenting observer” (Held & Thompson, 243-4). After his turn to the
communicative paradigm, Habermas barely mentions technology, apart from genetic manipulation
and neuroscience, which I discuss in Chapter 7. He treats technology as a largely value-neutral, norm-free tool of instrumental action.

Habermas is correct not to demonise modern technology or instrumental rationality as Horkheimer and Adorno had. The problem is that he neglects the fact that forms of technology unwittingly, yet unavoidably, express the cultural specificity of the values and norms of their designers, manufacturers, consumers and society as a whole. What characterises capitalist technological commodification is its focus on maximising exchange value, so suppressing other values and meanings found in creative relations to work, relationships or tradition. Feenberg argues that modern technology functions like a steering mechanism (Feenberg 1999, 166-73). It embodies materialised “norms” and “prescriptions” that guide our behaviour non-consensually, thereby exhibiting the same ambiguity that characterises money and power. Whilst it liberates us from risks and burdens of negotiation, it constrains our social possibilities. For example, buildings, their internal arrangements as much as their arrangement within an urban environment, both constrain and enable interactions – the manner in which, we stumble across, meet, acknowledge, see and be seen with, and by, and spoken to, and speak of, certain others. They do this not by the mere arrangement of physical form, but by a language of referring to socially given categories. Like Foucault’s power/knowledge, this is not something we necessarily need to (or even can) escape. Scientific-technological artefacts don’t merely “reflect” collective values but, like juridical categories, actively intervene in the lifeworld and mould us in ways of which we are barely conscious.

Whilst Foucault’s panopticon illustrates this well, we must not forget large scale configurations such as electricity grids, transport infrastructures and of course the internet. Since such artefacts interact with, produce, are produced by, enable, are enabled by, practices, concepts and juridificatory categories, we should in fact be talking about ensembles. Such ensembles also involve media like power and money, institutions and actors and so transcend the binary division between lifeworld and system, colonised and coloniser. As part of a project of normalisation that has produced us, they bear directly on individual identity, on what we appear to be “by nature”, yet remain largely concealed in a background operating behind our backs. It is not clear that such interactive matrices could fit within the lifeworld/system schema without putting pressure on Habermas’ commitment to the power / validity distinction and the context-transcendence of validity claims.

122 See Chapter 5 for discussion of expressive dimensions of technology.
123 Chapter 7 discusses Habermas’ notion of context-transcendent truth.
Thus far, we see Habermas setting up a comprehensive theoretical framework consisting of categories derived from systems theory, the notion of the lifeworld, formal pragmatics and developmental and evolutionary studies. Intended to bear critically on contemporary society, this theoretical structure should be seen in relation to the previous generation of Frankfurt theorists. Like his predecessors, Habermas seeks to undermine the pseudo-objectivity of social processes and theories which support them. Unlike them, he systematically integrates sciences, to provide a perspective external to lifeworld intuitions to build a comprehensive social theory with a robust normative basis. He reveals the presuppositions necessary for communication which, whilst fallible, provide as much certainty as possible to orient judgement and action. Unlike his predecessors, he thinks that emancipation does not entail rejection of the cultural legacy of Enlightenment but rather its completion. The lifeworld, threatened as it is, still contains normative potentials that are not exhausted. The problem he sees is that the dimensions of rationality available to modernity have not been actualised and institutionalised in a balanced way. I want to conclude this chapter by stepping back from the detail of Habermas’ account to examine the theoretical structures he has put in place. Whilst these structures capture the sense in which systems come to autonomously restructure lived experience according to their logic, Habermas doesn’t sufficiently elucidate the complex nature of the system/lifeworld relationship.

Critics have questioned whether Habermas has captured the universal and unavoidable presuppositions of reaching understanding in language by his distinctions between different “worlds”, along with the corresponding validity claims, attitudes and value spheres. McCarthy notes that such structures are to be found only in certain cultures at certain times (McCarthy 1991, 135). Searle argues that Habermas’ communicative action is not a universal basis of communication because societies have existed which have not been oriented towards mutual understanding (Searle in Muller-Dohm, 472 n.4). Habermas, responds to these concerns with a Hegelian move. Communication has a history and evolves. “This does not mean that such cognitive structures appear all at once, whether in ontogenesis or in social evolution.” (cited in McCarthy 1991, 135) The universal significance of these structures is shown by their mastery which represents the developmental-logical unfolding of species-wide competences, revealing unconstrained agreement in language as a fundamental ideal for social co-operation.

Viewed in terms of this developmental perspective, Habermas’ claims of universality appear more plausible. His developmental account draws on the authoritative (though fallible) backing of science,
to account for changes which he can interpret directionally, as progress.\textsuperscript{124} However, Habermas’ characterisations of progressive rationalisations are inevitably drawn from a Western, modern perspective and hence assume their progressive nature. Lacking an adequate grasp of developmental goals in other cultures, we fall back on our own dominant, though parochial, criteria which we project as universal. McCarthy argues that such cognitive-developmental paradigms are open to charges of ethnocentrism, scientism and rationalistic bias (McCarthy 1991, 134-9). This does not suggest to me that we should necessarily reject differentiation as an achievement of cognitive development. We must however, be wary of authoritarianism which \textit{insists} that the differentiation of domains of reality and corresponding validity claims are the ultimate end-points to which rational development is directed. As McCarthy suggests, there is much we could learn from pre-modern cultures, not only the things we have forgotten or repressed, but “something about how we might put our fragmented world back together again. It is not a matter of regression but of dialogue” (McCarthy 1991, 151).

However, we must grant that \textit{some} type of functional perspective is required to reconstruct a critical emancipatory project. We saw Foucault’s project gain traction by raising the question not what power is, but what power does, its function. Habermas is correct to insist that a dual perspective is required for critical social theory. Whilst the lifeworld concept gets at processes of symbolic reproduction, a functional “scientific” account brings into view social processes that occur “behind our backs”. Whilst McCarthy thinks Habermas has given away too much to systems theory, he agrees that “the idea that history has not been made with will and consciousness, that unplanned consequences, unrecognised interdependencies, uncomprehended systemic dynamics hold sway over our lives like a second nature, is quite naturally spelt out by [a functional account]” (McCarthy 1991, 177). The sort of functionalist perspective that is required has already been articulated by Habermas’s earlier critiques of systems theory in which he argued that functionalist social theory is not a form of empirical-analytic inquiry. The “needs” of social systems cannot be empirically established since they are not directly observed, but culturally mediated. A society’s “survival” is not a matter of objectively specifiable parameters.

A form of functionalism is required to provide an “interpretive framework”, although not on the model of biology or cybernetics. McCarthy cites Habermas’ call in \textit{On the Logic of the Social Sciences} for a functional analysis “that is hermeneutically enlightened and historically oriented and has as its aim not general theories in the sense of a strict empirical science, but a general interpretation of the kind we examined in the case of psychoanalysis …” This “historically oriented functionalism … “retains a hypothetical moment” (LSS 187-9). This earlier proposal suggests a type of hermeneutically informed functionalism that reveals hidden power structures without itself appearing

\textsuperscript{124} As we have seen, this is quite different from Foucault’s genealogies in which history provides the external perspective but he remains agnostic about progress.
to stand outside power. Social analysis would be undertaken from the standpoint of realising, to whatever extent possible, a social organisation based on unrestricted and undistorted communication.

Notwithstanding his tendency to treat rational reconstructions as “pure” forms of inquiry, Habermas’ theoretical framework is able to bring modernity’s pathologies into sharp relief (PKHI, 182). His social theory bears on, for example, the dysfunctional state of modern democracies by concepts such as colonisation of the lifeworld, fragmentation of consciousness and systematically distorted communication. His categories enable analysis, diagnosis, refinement, modelling and questioning of society from a standpoint outside the everyday lifeworld. In response to criticism, Habermas can stress the falliblistic, formal and procedural character of his theoretical structures. His idea of a society based on agreements, arrived at in free and equal exchanges, is not the depiction of a concrete utopia. The pragmatic force of this context-transcending ideal is that its regulative function opens reason to self-correction, enabling a foothold for criticism of what is found to fall short of idealised presuppositions.

Whilst providing an explanatorily fruitful diagnosis, there are still reasons to conclude that Habermas’ framework doesn’t address the complexity of the relationship between lifeworld and systems, particularly the role that science plays in that relationship. To reiterate three points: Firstly, in terms of Habermas’ theory, scientific discourses could be considered to operate as links that formulate, constitute, describe, interrogate and justify systemic power to the lifeworld. We could say that scientific discourses, split-off from the lifeworld, derive an “aura” of authority from the faith that their claims can be “ultimately” redeemed and their authority held to account within a broader validity realm.125 This authority serves to maintain social cohesion, avoid conflict and generate consensus. Habermas’ account needs to be expanded by examining the role of science in terms of this functional role within the juridification processes through which colonisation operates, since these processes explicitly adopt and apply elements of scientific and quasi-scientific discourses.

Secondly, a cost of Habermas theoretical structure is the difficulty in analysing forms of oppression in terms of socialisation within the lifeworld. Given that he primarily associates power with systems, power in the lifeworld seems insufficiently accounted for. The operation of power within the lifeworld is never developed into the sort of model of subjection found in Foucault’s power-knowledge. The possibility of a counter-movement from lifeworld to systems does not tell us anything about the domination which Fraser discusses and that appears endemic to the lifeworld. Habermas may respond

125 I mean “aura” in Benjamin’s sense of that which has authority derived from its appearance of naturalness, legitimating and inducing a passive affirmation of the status quo. (see Benjamin 1969)
that he acknowledges power in the lifeworld, but in the rationalised lifeworld of modernity, the
overwhelming threat is to the lifeworld’s communicative capacities from autonomous systems
detached from social reflection and accountability.

Finally, I have discussed how Habermas’ framework implies that scientific technology is “neutral”.
Feenberg is correct to suggest technology is not neutral, but a steering mechanism (Feenberg 1999,
166-73). It does not simply solve problems in the most efficient way, but in ways that particular
societies recognise and articulate what counts as “efficient”. Modern technology unwittingly, yet
unavoidably, embodies particular “norms” and “prescriptions” that steer behaviour non-consensually,
in much the same way as media such as money or power. The relationship between the perspectives
of lifeworld and system and the role of science in this relationship requires a more nuanced account
than Habermas’ colonisation thesis provides, even with the softening of his original thesis to
accommodate a two-way movement. I will return to the Habermas’ theoretical structures in Chapter 8,
where I will show that both Habermas and Foucault are unavoidably restricted by stances on which
their analyses nevertheless depend. In the next chapter, I will turn to Habermas’ discussion of facts
and values and his shift to a reference theory of truth. We will also discuss the public realm and,
looking at some specific examples of contentious scientific innovations, philosophy and religion.
Chapter 7
From the Theory of Communicative Action to the present

In the latest phase of Habermas’ work, from TCA to the present, we see the ongoing development of themes that had defined his earlier work – democracy, communicative reason and non-reductive naturalism accommodating lifeworld intuitions. During this period, he maintained his frequent activities as a public intellectual. Sharing Foucault’s suspicion of the “universal intellectual”, “the mandarin with his great gestures [who] enjoys an uncritically granted advance of trust”, Habermas brings his philosophical/scientific work to bear on specific issues of public concern in a way that demonstrates the power of communicative reason in shaping political culture (Muller-Dohm, 252). Just some of the issues which he has spoken out on - German re-unification, post-national Europe, the “historian’s debate” about German guilt, the Gulf war, immigration, multiculturalism, religious fundamentalism and the GFC – indicate his ongoing commitment to the public sphere (Muller-Dohm, 265, 277, 265).

I will firstly turn to Habermas’ major work Between Facts and Norms (hereafter BNF), in which we see the theoretical counterpart to this activity, oriented by an ideal of democracy as the free circulation of opinions and unrestricted interchange between specialised knowledge and everyday life. Here Habermas reworks his ideas on the public sphere to incorporate communicative theory into a theory of deliberative democracy. I will then consider, in Truth and Justification (hereafter TJ) Habermas’ reformulation of his 1960’s theory of truth. I will also discuss his “soft” naturalism, “weak” transcendentalism and the entanglement of facts and values. I will next consider the role and relevance of philosophy in the contemporary post-metaphysical context, before turning to a case study of Habermas view of science’s relation to society. In The Future of Human Nature (hereafter FHN), Habermas questions genetic intervention, which he sees as threatening the lifeworld’s categorical distinctions. I will then tease out his understanding of the place of religion in post-secular societies. I will conclude with a further case study on his discussion of free will and some concluding remarks.

1 BETWEEN FACTS AND NORMS.

In BFN, Habermas describes the processes by which public opinions are ideally translated into binding decisions in modern democracies. This includes addressing contentious scientific and technological proposals which challenge existing laws, rights or values. Decisions about these proposals – for example, genetic engineering or nuclear energy - require arguments for and against to
be openly considered and eventually converge as binding decisions. As I will discuss, this is only one way in which science is drawn into society.

Equipped with substantial theories of modernity, society and communicative action, BFN returns to the concerns of the public sphere that Habermas had focussed on 30 years earlier. In STPS, he had argued that the bourgeois public sphere of the 18th century had been taken over by organised capitalism, with decisions increasingly legitimated by means of mass persuasion by profit-oriented mass-media that trivialised democratic aspirations. BFN offers a more detailed and differentiated account the public sphere, its diversity and the system imperatives which impinge upon it. Any theory of democracy must incorporate these complexities to understand sovereignty as, not a united people in possession of one will, but intersubjectively, as the formulation of public opinion by reasoned argument, within both political sphere and the informal public sphere (BFN 301).

BFN marks a distance from TCA, in particular its colonisation thesis which had emphasised the systemic obstacles to the lifeworld’s influence on the political system. Habermas now focuses on the requirements for a public sphere in a genuinely democratic society. His democratic theory is an interpretive account of actual practices, from which he reconstructs the norms presupposed by political discourse. Habermas is guided by the conviction that law and morality are connected, despite having separated into independent spheres in modernity. Legal norms require moral legitimation which they acquire from their free and rational acceptance by citizens (BFN, 408). People see the law as just by virtue of the deliberative procedures by which it is instituted.

Habermas situates his theory of deliberative democracy as a response to the dominant traditions of liberalism and republicanism. He sees modern autonomy expressed politically as the freedom that citizens have in participating in the making of laws which they, at the same time, must live under. This results in two forms of autonomy being coded as rights - the mutually supporting and interlocking republican and liberal notions of public autonomy to participate, and of private autonomy as subjects with equal freedom against unjust interference (BFN, 408). However, to not pre-empt public deliberation, Habermas rejects attempts to prescribe concrete rights beyond their abstract formulation. He derives very general norms that regulate communicative action (BNF, 127). We encountered these general communicative norms of equality, openness, inclusion, generalisability in STPS (STPS, 36). These are not preconditions for democracy, but counterfactual presuppositions which provide ideal standards for deliberation, against which actual deliberation can be criticised (BFN 322-3). In following these norms, democratic deliberation is “the result of constructive opinion and will-formation” (BFN, 336-7). In the course of deliberations, in which the fully prescriptive meaning of rights is worked out, the initial interests of citizens and definitions of problems may be transformed.
To articulate the nature and role of public opinion in relation to decision-making, Habermas introduces a two-track model by distinguishing informal from formal discourses. The informal track revolves around the public sphere – the network of communication flows that extend from face-to-face egalitarian discussion to impersonal mass-media communications. These streams of communication are “filtered and synthesised in such a way that they coalesce into bundles of topically specified public opinions” (BFN 360). Informal public opinion formation is inclusive and critical and is primarily oriented towards the discovery and definition of problems. Because it is diffuse, unregulated, open and chaotic, it cannot coalesce into binding decisions. The formal track comprises discourses within parliaments, government bureaucracies and courts, in which formal fairness, civility and equality is ensured by procedural regulation. Discourse in the formal track is oriented towards arguments that issue in decisions. Only the legislature is capable of working through conflicts to reach binding decisions. Only judges and administrators can act decisively on laws. This procedural orderliness comes at a cost. Public opinion is filtered to fit the agendas of political parties or bureaucracies. By inhibiting the spontaneous and inclusive exchanges of opinions that characterise public opinion formation, the formal track reduces the quantity and quality of opinions that get heard. (BFN, 354-62).

To be legitimate, binding decisions must be steered by communications from the periphery passing through democratic and constitutional procedures, to the parliament and courts (BFN, 356). If civil society is structured to allow active and equal participation in political life, it can “ferret out, identify, effectively thematise latent problems of social integration which require political solutions.” A sensitive and active periphery can detect and identify new problems which it can “introduce via the parliamentary or judicial sluices into the political system in a way that disrupts the latter’s routines” (BFN, 358). Habermas notes that the nuclear arms race, genetic engineering, ecological threats, third-world poverty were all issues that weren’t initially brought up by exponents of the state apparatus, large organisations, or functional systems (BFN, 381).

Despite Habermas’ well documented doubts, critics have suggested that his account fails to raise concerns over the state of modern democracies, evidenced in cynical media manipulation, apathetic citizenry and abuse of social power. His account is thought to have abandoned the critical impulses of his earlier work. Scheuerman suggests that it is ambiguous between an ambitious proposal for radical democracy and a description of a defensive model in which democratic institutions exercise an attenuated check on market and administration excesses (Dews 1999, 153-178). Lyotard argues that in the transmission of communicative power, what starts as discourses about injustice is transformed into bureaucratic discourses about balancing interests and techno-scientific discourses about economic growth (Lyotard 1988, 157). Cook portrays Habermas affirming the status quo, “uncritically
espous[ing] the norms and values currently endorsed in the constitutions and institutions of contemporary liberal democratic states” (Cook 2001, 95).

These substantially unfair criticisms call for clarification. Habermas wouldn’t deny the internal threats that actual democracies face. However, he would insist that the norms of deliberative democracy can be still discerned within actual democratic-legal discourses and that these general communicative norms provide a standard against which actual practices can be judged. Habermas wouldn’t deny the strong connections between government, bureaucracies and technical elites. However, he would argue that specialised subsystems are not entirely self-contained. I will now consider the dissemination of science and technology into democratic societies, roughly in terms of Habermas’ framework.

Since science is not homogeneous or monolithic, it can operate in various ways within Habermas’ model, with different scientific discourses often playing different, or even opposing, roles in relation to single issues. I have already mentioned the justificatory role that science plays, specifically in the formulation of legal codes and assessment, in terms of those codes, of proposals such as for nuclear energy or genetic intervention. Whereas within the informal circuit, subjective, emotional and religious views have their place alongside science, within the formal circuit of decision-making, science is granted an authoritative role. When presented to judges, politicians or administrators, scientific discourses are given primary weight in assessing potential impacts.\(^{126}\) As discussed in Chapter 6, formal decisions are guided and justified by scientific studies in which meanings, significances and values are straight-jacketed into methodologies digestible by decision-making authorities. Quantitative approaches enable values, where acknowledged, to be weighed against other “social goods”, resulting in decisions that appear robustly defensible. The abstractions and quantifications required to filter out subjective and qualitative aspects are perhaps an inevitable outcome of the accountability that democratic decision-making demands. Nonetheless, by the use of science as the most authoritative arbiter of validity in legal contexts, the lifeworld’s interpretative capacities are substantially mediated.

Scientific discourses also operate less explicitly. As background knowledge, they align themselves with legal paradigms that tacitly guide consideration of social issues such as multi-culturalism, same-sex relations or women’s rights. Habermas refers to the judiciary’s tacit reliance on two legal paradigms representing conflicting views of society. The liberal paradigm is based on the idea of an

\(^{126}\) By “scientific discourses”, I refer to cost benefit analyses, environmental impact statements, social impact statements and so forth which draw on a range of scientific disciplines and tend to be framed exclusively in objectivistic language.
association of independent entrepreneurs, exercising their negative freedoms, free from government interference. The 
"welfare" paradigm responds to economic inequality (which the liberal paradigm sees as just and natural) by holding the politically constructed economic system, rather than individuals, responsible. Habermas claims that growing social complexity and diversity has resulted in increasing conflict between paradigms, leading to the development of a reflective awareness of their limits. He argues that we need to link the development of concrete meanings of subjective rights to an ongoing process of democratic interpretation. This process harnesses a higher-order "proceduralist" paradigm to determine the most appropriate paradigm for particular situations, by requiring citizens to collectively discuss on a case-by-case basis the extent to which situations can be defined by either paradigm (BFN, 393).

As an example, Habermas discusses women’s rights. Both liberal and welfare paradigms, appealing respectively to gender difference-blind and difference-sensitive principles, are vulnerable to reliance on false stereotypes of what is biologically determined. In what Foucault refers to as “normalisation”, both paradigms employ overgeneralised classifications of gender identity which tend to reflect the status quo of contingent social power arrangements (BFN, 422-3). Habermas, like Foucault, is wary of biological essentialism (BFN, 425). Rather than appealing to science to find a truth or essence in human nature to define rights, Habermas’ proceduralist paradigm views women’s rights as undefined, emerging only through a struggle for recognition in the public sphere. Women themselves, as the affected group, must clarify the “relevant aspects” and standards that define equality and inequality in public discourses (BFN, 420, 425).

This proceduralism could be extended to scientific discourses. Given the adversarial nature of science within a social context, it is not surprising that certain scientific discourses spawn oppositional scientific counter-discourses. In this vein, Schnaiberg’s distinguishes “production science” from “impact science” (Gould et al). Production science disciplines might include typical applications of physics, chemistry and their technologically-oriented derivatives such as engineering, which are oriented towards industrial progress. Impact science might include ecology and areas within the human sciences which ask awkward questions about damage wrought by applications of production science. It is plausible that the adherents of these discourses are motivated by deeply-held pre-scientific attitudes, such as attitudes to risk and our place within nature as, for example, either one of nurturing stewardship or literal interpretation of the biblical injunction to dominate (Genesis 1:28).

My point is not merely that scientific discourses can be distorted by interests. Nor I am I saying that these discourses are irrational, but rather that they spring from pre-scientific commitments which themselves require articulation within public discourse, in a manner suggested by Habermas’ procedural approach to legal paradigms.
As well as conscious and deliberate use of scientific discourses in decision making, science permeates the lifeworld barely noticed. The scientific conjectures of, for example, neuroscience, genetics and evolution, profoundly influence the understanding of the self and society and their relationships. Technologies which are incrementally but continuously being refined and enhanced, barely noticed unless they cross the threshold of laws, conventional norms or values, mould us by the values materially embedded within them. Ever-resourceful capitalism invents new disciplinary technologies aided by new management sciences. This moulding of the lifeworld by science isn’t captured by either TCA’s colonisation thesis or BFN’s procedural democracy because in neither work, does Habermas venture from the idealised realm of language and communication to lay out, as Foucault does, a genealogy that includes things and bodies. Given the level of generality of his work, it is difficult for Habermas to theorise either the pre-scientific commitments behind science or the barely noticed material penetration of technology into the lifeworld. However, we will see Habermas suggesting that when scientific-technological innovations threaten the lifeworld, philosophy is available as an interpreter. Before coming to philosophy’s role, we must firstly update Habermas’ view on truth, to which science is oriented.

2 TRUTH AND JUSTIFICATION

Whilst relying on fundamental concepts of objectivity, reference and validity, Habermas provided no theoretical basis for these concepts in his communicative theory. His “consensus theory” had sought to explain truth by appealing to “ideal” conditions of justification. Claims are true if they can be argumentatively redeemed in discourses, that lead to a non-coercive and justificatory consensus. But this came too close to identifying truth with ideal justification. To address this problem, Habermas returns to a line of thought he had put aside since KHI (TJ,1).

Truth and Reference

Habermas has consistently rejected reductive naturalism which explains knowledge by an independently existing nature causing our ideas. But without unmediated access to reality, we have no knowledge of a world existing independently of our descriptions. Rorty suggests that we do away with such realist intuitions, deflate truth and replace “objectivity” with “solidarity” (Rorty 1981, 21-34). Habermas disagrees. Statements are true not because the linguistic community agrees, but because the states of affairs they describe obtain (although that can only be established by means of other statements). He thinks Rorty’s radical contextualism assimilates truth to holding-to-be-true, and can only lead to a self-contained circle of scepticism and relativism. This assimilation loses the “the sense in which statements referring to objects of an objective world that exists independently of its description are true – a sense which transcends the justification for these statements” (cited in Muller-
Dohm, 372). Validity claims must be defendable “in all possible contexts, that is, at any time and against anybody” because by its very nature, argumentation points beyond all particular forms of life (OPC, 367). This constitutes a moment of unconditionality, a moment which enables “the process of justification [which] can be guided by a notion of truth that transcends justification” (OPC, 372).

Whilst requiring that truth transcend any justificatory context, Habermas refrains from defining truth. He adopts the pragmatist approach of directing our attention to the role that truth plays and how it functions. Habermas’ Kantian pragmatism holds that although the world is independent of the human mind, the structuring of the world into kinds, individuals and categories is a function of the mind. The belief in a shared objective world is transcendentally presupposed, but only in the weaker sense that any attempt to deny it could only be put forward against the background assumption of a shared objective world (BRN 41). “Weak transcendentalism” is balanced by “soft naturalism”, which presumes that language can refer to a language-independent reality. Language doesn’t fully determine what we can know of the world, since by engaging with this independent reality we can revise the meanings of terms (TJ, 26). Every action is an attempt at successful coping, provoking learning by engaging with either the causal resistance of nature or the normative resistance of others. Reality constrains our practices in tangible ways which provides a basis for both a robust notion of objectivity and learning to successfully cope with the resistance of the world or others who criticise our claims (TJ, 16).

Truth can only operate as part of a learning process if revised beliefs still refer to the *same objects*. Scientific revolutions in which fundamental descriptions are revised render such reference paradoxical. It is not clear whether we have a better understanding of the same thing, or are talking about something different. Habermas wants to support an account of scientific progress by drawing on the formal pragmatic presupposition - that speakers *refer* to the same things – which is independent of the specific, and possibly different, *descriptions* that speakers might associate with a referent (TJ, 33). Habermas appeals to Putnam, for whom the scientific essence initially fixes the role of descriptions by rigidly designating a set of conditions considered normative by the community.¹²⁷ However, within the context of ongoing scientific debate and experimentation, scientific definitions operate like revisable operational definitions employed by ordinary persons. Rather than imposing a fixed set of necessary conditions, operational definitions enable us to speak of progress by referring to the same things.

¹²⁷ A “rigid designator” designates the same object in all possible worlds (Kripke 1980).
**Action and Discourse**

The distinctive feature of Habermas mature theory is that “Janus-faced” truth stands between discourse and action, relating these two domains to each other. On one hand, truth is operationally effective in action as an unthematised certainty (OPC, 364). We hold beliefs to be unconditionally true as long as they provide practical certainty within contexts of action (OPC, 372). On the other hand, once a belief ceases to be a successful guide to action, it becomes questionable and is drawn into discourse in which its truth becomes a function of justification. Truth operates here like Kant’s regulative ideals, to guide inquiry by transcending all justification, reminding us that no actual justification is sufficient and urging further and better reasons (OPC, 372). Rational discourse forces those participating in it to continue decentering their cognitive perspectives (TJ, 38). Bernstein objects to this sharp distinction between action and discourse (Bernstein 2010, Ch 8). Action is not completely unreflective and can involve deliberation. Discourse frequently involves actions, experiments and interventions. Action and discourse are better seen in terms of degrees along a continuum between completely routinised unreflective behaviour and discursive reflection disengaged from action and its consequences.

I agree with Bernstein’s objections. The difference between the role of truth in contexts of discourse and action seems to be that in discourse, a thematised belief is consciously and explicitly held as hypothetical in a sustained manner. In contexts of action however, a belief may be doubted yet remain sufficiently unproblematic, in terms of the likelihood and possible consequences of its falsity, for it to be set aside from deliberation. In action, beliefs are treated “as if true” in order that we can stop deliberating and get on with our lives. The difference between the operations of truth in these contexts is a difference in degree. If Habermas’ sharp distinction between discourse and action can’t be sustained, Bernstein thinks that it isn’t necessary to appeal to “justification-transcendent truth” since it has no explanatory function. Rorty agrees. “The whole pragmatic force of the claim that truth is not conditional is to express willingness to change one’s mind if circumstances alter, not to explain or justify this willingness” (Rorty 2000, 57). But it is not clear that Habermas is attempting “to explain or justify” willingness to change one’s mind. Habermas is elucidating the concept “truth” in terms of the world being as it is, independently of our beliefs about it. To say that truth is “justification-transcendent”, seems a way to give sense to the fallibility of our beliefs in the light of this independent world which is only revealed partially by our actions and never captured fully by our beliefs. Without a concept of truth that transcends justification, we are caught in a performative contradiction (OPC, 357). Habermas cites Wittgenstein: “If you tried to doubt everything, you would not get as far as doubting anything. The game of doubting itself requires certainty” (OPC, 356). Whilst we can adopt the observer’s perspective towards our epistemic commitments, thus suspending their force, we can only do so from a participant’s perspective of unquestioned certainty. “Subjects engaged in their practices refer to something in the objective world, which they suppose as existing independently and
Both distinctions - action/discourse and participant/observer – are untenable as rigid dichotomies. Discourse can be understood as a form of action, whilst action can be a form of internal discourse not fully conscious. However, this blurring of distinctions doesn’t undermine the role of the notion of unconditional truth with which Habermas attempts to do justice to his realist intuitions. There is only one way the objective world is, and that is not up to us. Habermas articulates his theory of truth by bringing this notion into relation with the notion of truth that exceeds all justification.

**The Entanglement of Facts and Values.**

Whilst truth refers to an objective world, Habermas thinks that moral claims lack such reference and are solely directed towards ideal consensus. When a moral norm is contested, we seek justification by a consensus of all those affected under conditions of an ideal speech situation. The agreement reached constrains our behaviour by projecting an ideally inclusive community of all speakers. A fundamental feature of Habermas’ Kantian pragmatism is this sharp distinction between practical and theoretical reason (TJ, 248). Putnam argues that Habermas’ distinction is too sharp. Science assumes epistemic values such as coherence, plausibility and simplicity in selecting between theories. Whilst distinct from moral or ethical values, these values are still judgments of “what ought to be” in the case of reasoning (Putnam 2002, 31). Non-epistemic values are also entangled with facts. To call someone “cruel” is both to criticise and describe (Putnam 2002, 39-40). Facts carry an evaluative dimension and values entail facts (cf Murdoch 1970). In determining moral norms, learning processes frequently involve coming to a better understanding of factual consequences and implications (Bernstein 2010, 196).

Whilst Habermas acknowledges this entanglement, his strong realist intuitions lead him to distinguish facts from values and criticise Putnam’s notion of “ought-implying facts” (TJ, 224). What justifies values is their reference to a counterfactual ideal agreement (TJ, 230). Habermas grants an analogy between truth and moral norms but this is not to identify them (TJ, 238). “Truth” is a notion that applies to the independent world of facts only. Moral judgements apply to the social world and are subject to a different kind of validity. I am sympathetic to Habermas’ “realist intuitions”. Stubbing one’s toe demonstrates the resistance of an objective world of a different kind to the resistance of others’ disagreement. There is also a difference in how facts and values fit with the world. “Empirical judgments say how things are in the objective world whereas evaluative judgements enjoin us to value or treat something in our lifeworld in some way or other” (TJ, 224). However, Habermas’s sharp distinction neglects our activity in constituting concepts and facts, the sort of constitution Foucault
sought to highlight and which suggests that what is “up to us” is a *matter of degree*. In similar vein, Dewey sees facts as “not complete and self-sufficient in themselves. They are selected and described … for a purpose, namely statement of the problem involved in such a way that its material both indicates a meaning relevant to resolution of the difficulty and serves to test its worth and validity … They are not merely *results* of operations of observation.” (Dewey cited in Bernstein, 199).

Habermas may accept this way of blurring of the fact/value distinction. If so, the question would then be whether we can still consider this distinction to have a universal basis, or whether it is merely a local historical artifact, as Foucault might suggest. I am unable to further analyse, or attempt to adjudicate the differences between Habermas, Putnam and Bernstein here. However, what is relevant, and something Habermas recognises, is that science explicitly thematises the validity claim of truth, whilst repressing the other validity dimensions, which retreat into the background of the lifeworld. To ignore these dimensions and so neglect the unavoidable entanglement of scientific facts with values implied by their historical constitution, selection, context of application and technical potential, is to misrecognise science’s practical implications.

3 THEORY AND PRACTICE REVISITED.

Whilst Habermas emphasises the sharp distinction between validity spheres, he champions philosophy’s role in integrating the truths of science into the broader validity basis of the lifeworld’s “unity of reason”. He formulates his argument by firstly, situating philosophy in the historic context of a series of shifts and re-alignments in the relationship between theory and practice (TJ, Ch 7). By showing that philosophy does not fulfil the same continuous role or pursue the same eternal questions, Habermas teases out what philosophy should be in the present. One thing that it shouldn’t be is found in Habermas’ critiques of positivism and reductive naturalism. The positivistic self-understanding of the sciences is not scientific, but metaphysical. Both positivism and metaphysics are committed to a theoretical attitude that seeks freedom from the distortions of interests to describe the universe, just as it is, in its law-like order (KHI, 303). In a post-metaphysical world, Habermas insists that we can no longer hold to notions of completely detached observers of an independent world.

Habermas traces the transformation of the ancient ideal of contemplation of “objective reason” into Kant’s epistemology which becomes foundational in the form of insights into the *a priori* ahistorical transcendental conditions of different types of experience. By this foundational insight, philosophy becomes an “usher”, able to show the sciences their proper place, and a “judge”, setting limits on their claims in relation to other areas of culture (MCCA, 2). In contrast, Habermas wants philosophy to be critically directed to its subject matter whilst reflexively aware of *its own* embeddedness in an
intersubjective lifeworld. Rather than informed by metaphysical conceptions of the whole of history or nature, or the imagined aspirations of a collective subject, theory should be seen as fallibly informing practice whilst, at the same time, being modified and corrected.

Habermas turns to pragmatism and hermeneutics which embrace the idea of cognition mediated by language and linked to action in a web of everyday communications and life activities. He is critical when these approaches contrast their “narrow objectivistic conception of science” to philosophy which they see as “illuminating or awakening instead of being objective.” Science is completely distinct from philosophy, which they see as making no claim to validity (MCCA, 12-13). Rorty varies this theme by his distinction between “normal discourses”, which have reliable criteria for settling disputes and make progress, and “abnormal discourses” which are incommensurable when their basic orientations are contested. Rather than dying out or morphing into normal discourse, some abnormal discourses persist as “interesting and fruitful disagreement” sufficient to themselves. According to Rorty, philosophy verges on this “edifying” condition, once it has got beyond any pretentions to problem solving. Whilst sympathetic to Rorty’s view that “philosophy has no business playing the part of highest arbiter in matters of science or culture”, Habermas rejects the idea that philosophy can be “outside the sciences, without being immediately drawn back into argumentation, that is, justificatory discourse” (MCCA, 14). He is uncomfortable with Rorty’s claim that philosophy, unlike science is merely edifying. I share his discomfort. Philosophy entails holding to rational commitments that connect practically to the social and objective worlds that science has in view.

Habermas rejects any sort of reductionism, for example, the lifeworld to a system, reasons to causes or freewill to brain states. By keeping both poles of such dichotomies in play, philosophy can criticise both reductive naturalism and metaphysics. In so doing, philosophy must respond to “the particular need of modernity, which is, after all, bereft of any guidance by models of the past” (TJ, 284). Modernity has broken with the continuity of traditional models of behaviour and reasoning and must develop its own normative understandings. Philosophy facilitates this by a role we saw in Foucault’s analysis of Kant’s late occasional writings. Against the “highly specialised business of a remote discipline”, Habermas champions “exoteric approaches” which “have public impact because they face up to the problems that confront philosophy from both private and public life” (TJ, 284). With this modern fallibilist sensibility of a discipline aware of its situatedness in the world, “philosophy is forced to drop the claim of holding the key to Truth” and finds its place in functionally differentiated roles within the modern world (TJ, 285).

However, “philosophy cannot completely immerse itself in any one of its social roles”. It can only fulfil these roles by “at the same time transcending [them]”. To fully immerse itself would “rob it of its best … a kind of untamed thinking that is neither channelled nor fixed by any particular method”
(TJ, 286). Philosophy, as an interpretive practice, cannot be overly prescribed by method. Science however, must be guided by methods which rule out in advance certain types of objects and claims as unscientific. To “completely immerse itself” in, say the role of elucidating science, philosophy would have to conform to science’s strictures - the idealisations, abstractions and limitations that scientific method necessarily requires. However, by “at the same time transcending [this role]”, philosophy can have these strictures in view, whilst retaining its connection to the lifeworld.

Since the 17th century, science has become increasingly independent of philosophy. Today fallibilist philosophy, stripped of foundationalist claims, “can do no more than respond to the independent developments of sciences that have become autonomous” (TJ, 286). Philosophy can also enter into co-operation with the sciences. Rather than a clear-cut division of labour between philosophy and science, Habermas refers to the blend found in Marxism and psychoanalysis, which is typical of many social sciences by virtue of containing “a genuine philosophical idea”. Habermas cites Freud (symptom formation through repression), Durkheim (creation of solidarity through the sacred), Mead (identity formation function of role taking), Weber (modernisation as the rationalisation of society) and Chomsky (language acquisition as hypothesis testing) as paradigmatic examples of theories which contain embryonic philosophical ideas which at the same time pose empirical and universal questions (MCCA, 15).

These sciences are not “immature”, eventually converging as a unified science, “the triumphal march toward objectivist approaches such as neurophysiology,” but rather stages “on the road to the philosophization of the sciences of man” (MCCA, 15). Here philosophy operates as a “stand-in” for “empirical theories with strong universal claims, the primary candidates being “reconstructive sciences” which presume to explain the universal bases of rational experience, judgement, action and linguistic communication by anonymous rule systems that competent subjects follow” (TJ, 287). Rather than philosophy pretending to the roles of usher or judge, it collaborates with science “by an auspicious matching of different theoretical fragments.” 128

Whilst defending the rationalisation by which different validity spheres became autonomous, Habermas wants to establish a balance between the separated moments of reason in the lifeworld. As we saw, the increasing specialisation of science has led to its remoteness from the lifeworld. Whilst Foucault sought to excavate the common ground of the episteme, in which science and non-science overlap and influence each other, Habermas casts their separation as a progressive rational development. But now, split-off elitist spheres fragment consciousness and weaken the lifeworld’s communicative resources which offer the only real alternative to exerting influence in more or less

128 Examples of collaboration between philosophy and empirical investigations, include cognitivist ethics, psychology of moral development, speech act theory and empirical approaches to the pragmatics of language.
coercive ways (MCCA, 19). To co-ordinate actions without resort to coercion, we must draw on “a cultural tradition that ranges across the whole spectrum, not just the fruits of science and technology” (MCCA, 18). Here philosophy takes on “the role of interpreter on behalf of the lifeworld … refurbish[ing] its link with the totality” (MCCA, 18-19). Philosophy is able to “preserve unity across all the disparate aspects of validity” without reducing or levelling them (TJ, 287). It can initiate an interplay between various dimensions of validity that have “come to a standstill like a tangled mobile” (MCCA, 19). Philosophy should drop its role of arbiter, inspecting and judging aspects of culture, and act as interpreter, unravelling these tangled parts and connecting them into the more comprehensive whole of the lifeworld. By thinking in terms of the whole, philosophy is able to develop interpretations that are normatively charged with practical intent. By its close relation to both sciences and to common sense, it “can criticise the colonisation of a lifeworld that has been gutted by trends of commercialisation, bureaucratisation, and legalism as well as scientism” (TJ, 290). A dialectic between potential and will can only be established within the shared lifeworld in which reasons can connect, back and forth across validity spheres, from descriptions of what is, or what could be, to prescriptions of what should be.

Philosophy must maintain its “interface” with science or “it would lose the very insights of its own that it needs in order to fulfil its exoteric roles” (TJ, 287). Science can challenge philosophical commitments by demanding that philosophy make its positions perspicuous, which it must do in the light of both scientific claims which must be taken seriously and common sense intuitions which can’t be abandoned. As we will see, this exercise can also serve as a corrective to incautious claims by scientists and their uptake by a naïve laity. Philosophy can undertake a “diagnosis of our time in terms of which modern societies come to understand themselves” conducted “primarily in the philosophical form of an auto-critique of reason” (TJ, 290). Habermas is referring to modernity’s self-reflexive attitude of distanciation by ongoing critique, a tendency to understand ourselves by adopting an “outside perspective”. Philosophy as critical theory enlists science as a partner to expose distortions of the lifeworld. Here science is not the ultimate authority, but rather a valuable perspective, though in need of translation and contextualisation within lifeworld commonsense. Macrosocial problems “are no longer visible from the perspective of closed self-referential functional systems such as institutionalised science and its discourses” (TJ, 289). Questions about the overall direction of science and technology should be considered within the “diffuse network of a public sphere anchored in civil society … [in which] highly complex societies become aware of significant failures and risks and can politically deal with [them]” (TJ, 289). In FHN, we see Habermas contribute to this public discussion, employing philosophy as an interpreter between science and the lifeworld. Typically he approaches such questions by initially seeking a prima facie intuition, for example, the deep-seated fears of cloning human beings. He identifies the rational core of such intuitions, providing the impetus for
reflections which provide an argumentative basis. The ensuing public discussions develop his position further.

4 THE FUTURE OF HUMAN NATURE.

In FHN, Habermas questions contentious applications of genetic science by considering the preconditions for our identity as autonomous moral agents. In the light of the technical possibilities for intervention in human nature, he thinks that central aspects of the normative orientation of the human species needed to be clarified. What is at stake is the “uncontrollability of the contingent process of human fertilisation that results from what is now an unforeseeable combination of two different sets of chromosomes” (FHN, 13). Habermas’ premise is that the randomness of the human genetic make-up is a precondition for the formation of individual identities as well as for the fundamental equality in relations between human beings. Manipulation of genetic material threatens the inviolable status of persons by treating them as mere objects. Genetic intervention potentially replaces the individual uniqueness and autonomy of a newborn with a new kind of inequality between humans.

Habermas’ critique is directed specifically against the idea of “liberal eugenics”, genetic interventions that manipulate traits in accordance with parental wishes and are freely accessible on unregulated markets. He doesn’t oppose biotechnological interventions per se, but rather wants to articulate the vague fears surrounding biotechnological programming. His analysis focuses on pre-implantation genetic diagnosis (PGD) which enables screening of embryos, to determine the presence of certain traits prior to implantation, thus opening the possibility of future genetic interventions for enhancement purposes. Habermas’ concern is that the biotechnology, not possible at present, enabling desired characteristics, may sever us from what is most human by a slippery slope passing from PGD to selective implantation, trait selection and cloning. By reducing persons to mere means, genetic enhancements have the potential to shape self-understanding in a way that threatens freedom and equality.

FHN can be seen within the larger context of the debate about “transhumanism”, in which Habermas is typically cast as a “bioconservative”, a label he happily accepts (DM 411). Transhumanism seeks to increase human health-span, extend intellectual and physical capacities, give increased control over mental states and moods by technologies like genetic enhancement, virtual reality, nanotechnology and artificial intelligence. Transhumanists promote the view that human enhancement technologies should be widely available, and that individuals should have broad discretion in their use (Bostrom 2007).
What human beings are “by nature” is increasingly coming within the reach of biotechnical intervention. From the perspective of science, this is just another frontier passed, in the ongoing extension of the ability to control and manipulate nature. Habermas sees normative regulation as generally adapting to technological development. Given the promises of gains in the scope of individual choice and prosperity, scientific research has generally been aligned with the basic commitment of liberalism which grants it its freedom. In fact, the history of medicine suggests a sceptical attitude to “moralising human nature”. Blood transfusions, vaccination, heart transplants, brain surgery, artificial insemination, artificial organs, have all been challenged but are now unquestioned life-enhancing measures.

However, given the systemic imperatives of economics linked to rapidly developing technology, Habermas thinks we risk unthinkingly endorsing a right to scientific research, only later finding technologies normalised prior to any informed consideration. We are, Habermas argues, being overwhelmed “by a lack of perspectives” (FNH, 18). “The development of biotechnology generates a dynamic which again and again overtakes the time-consuming processes through which society reaches a self-understanding about its moral aims” (Habermas cited in Muller-Dohm, 321). According to Habermas’ social evolutionary account, the archaic lifeworld adopted a number of fundamental categories - the made/the grown, the technical attitude/the clinical attitude, having/being a body, fate/autonomy - through which lifeworld members could grasp their world. He contends that these fundamental categories are now challenged by genetic technology. Driven by economic imperatives, such technologies may compound the erosion of the lifeworld’s communicative resources already taking place. FHN analyses the lifeworld in terms of this threatened categorical structure, showing how key distinctions are violated by certain genetic technologies.

Libertarians argue that parents have the right to an enhanced child and see concerns as just further dubious re-enchantments of inner nature opposing developments that will soon be accepted as normal. Against this, Habermas argues that the extension of technological control over our “inner” nature without our consent is distinguished by the fact that it “changes the overall structure of our moral experience.” Shifting the line between choice and chance affects our capacities to see ourselves as authors of our own lives and to recognise others as autonomous persons (FHN, 29). From this angle, Habermas’ attempt to ensure, against free-market eugenics, the contingency or naturalness of one’s creation can be seen as modernity’s self-reflective consideration of its limits.

**Grown and Made**

Liberal eugenicists argue that there isn’t much difference between the genetic enhancement of future individuals and socialisation. Just as a person’s dispositions and consequent life possibilities are moulded by socialisation, so they can be moulded by genetic interventions. Both occur without
express agreement. However, Habermas argues that socialisation proceeds by communicative action which for the parents, is connected to reasons. Even if this “space of reasons” is not yet open to the child, she still has the role of the second person in relation to expectations underlying the parents’ efforts. Adolescents still have the opportunity to critically reappraise restrictive socialisation processes. Genetic programming allows no such opportunity when it commits a person to a specific life project according to another’s intentions (FHN, 63). Habermas concludes that we may only have good reasons to assume the programmed person’s consent in the case of the prevention of clearly unacceptable diseases and disabilities.

Liberal eugenics affects the capacity for being oneself by establishing an unprecedented human relationship. The genetic programmer’s irreversible choice jeopardises a precondition for the moral self-understanding of autonomous actors, resting on the assumption that there is no definite obstacle to egalitarian interpersonal relations. We assume as a pragmatic ideal, that we could be in anyone else’s shoes. Whilst any genetic inheritance, programmed or natural, is a form of irreversible dependence, Habermas argues that, in the case of a natural inheritance, this only relates to a person’s existence, not their essence. Unlike couples who decide to have a child, the genetic programmer irrevocably breaches the reciprocity between persons by determining the essence of a future person. The person whose program has been deliberately fixed by another is, in principle, barred from exchanging roles with his designer (FHN, 65). This relationship is inimical to the reciprocal and symmetrical relations of mutual recognition required for a moral and legal community of persons, born free and equal.

The problem with liberal eugenics is that it presupposes the levelling out of the fundamental distinction constitutive of self-understanding - the distinction between what is manufactured and what has come to be by nature (FNH, 50). The self-emergent grown life that requires nurturing is distinguished from what a technician makes from inert matter, by its capacity for individuation and self-determination that exceeds causal intervention (FHN, 47). This distinction is reflected in fundamental attitudes of engagement with the world. A scientific agent combines the theoretical objectivating attitude of the disinterested observer with the technical attitude of an intervening actor to work instrumentally by imposing its purposes upon an object presupposed to be passive. However other attitudes, such as the clinical attitude, or attitudes of cultivating, healing or breeding “share a respect for the inherent dynamics of self-regulated nature” (FNH, 45). In these attitudes, practice adjusts itself to the object, and is responsive to its inherent purposes, potentials and meanings (FHN, 45).

This distinction still retains its power in the lifeworld that Habermas wants to defend.
Liberal eugenicists may attempt to assimilate therapeutic intervention with enhancement intervention. But in the case of therapeutic interventions, what matters is the clinical attitude of the therapist who in a performative attitude of a participant, anticipates the future consent, however virtual, of the person (FHN, 52). Where a person is genetically “healed” without his actual permission, virtual consent is assumed by a doctor who is not acting on an object, but relating to a future person. This differs from the case of someone who learns that his genetic makeup was programmed solely according to the preferences of a third person. This person may find it difficult to “own” the prenatal intervention as being theirs. This reflects the further distinction between having a body and being a body. The primary mode of experience is that of being a body (FHN, 49). Having a body is the result of a capacity to assume an objectivating attitude toward the prior fact of being a body. Habermas fears that upon becoming aware of being “made” by genetic manipulation, an adolescent’s participant perspective of being a body would “collide with the reifying perspective of a producer or a bricoleur” (FHN, 51). Where human nature is instrumentalised by genetic interventions, there is a conflict between the self-perception of the future person concerned and seeing that person’s nature as an “inner environment” to be acted upon.

**Natality, Freedom and the Limits of Eugenics**

Habermas suggests that the shift from the participant stance of living in one’s own body to the observer perspective which governed the intervention one’s body was subjected to before birth, is categorically different from the decentration of our geo-centric and anthropocentric worldviews by Copernicus and Darwin. The subjugation of our body to biotechnology involves the intention of another person intruding on our life history before our birth. This intrusion might erode the “capacity to be ourselves”, to speak with one’s own voice as “the person herself who is behind her intentions, initiatives, and aspirations” (FHN, 57). To be able to identify with one’s body requires that the body can be experienced as natural, “a continuation of the organic, self-generative life from which the person was born”. What is important is that “we experience our own freedom with reference to something which, by its very nature, is not at our disposal” (FNH, 58). Since human autonomy can only be understood in contrast with “fate” that limits freedom, the lifeworld makes the distinction between what we are, and what happens to us (FNH, 60). Fate is entwined with an origin that has its beginning beyond human disposal. To know oneself as the author of one’s actions and thoughts, one must be able to ascribe one’s origin to something beyond human manipulation, to a beginning, such as nature or God, which can’t be compromised by another person’s intentions.

Habermas turns to Arendt’s discussion of “natality” to elucidate this. Arendt claims that in acting, humans feel free to begin something new because birth itself, as a divide between nature and culture, marks a new beginning (Arendt cited in FHN, 59). Habermas interprets this to suggest that the significance of birth is that it marks the divide between the natural fate of the body and socialisation
that follows birth. It is only by maintaining the distinction between nature which is not at our disposal, and culture which is subject to historical contingency, “that the acting subject may proceed to the self-ascriptions without which he could not perceive himself as the initiator of his actions and aspirations” (FHN, 59). It seems that the very notion of “authentically being oneself”, as the source of one’s actions and thoughts involves, as a minimum, a natural history free from others’ intentions prior to entering socialisation at birth. Birth plays the role of a marker in the self’s history that gives the reference point of what one is, prior to any imposition of others’ intentions, and it is this reference point that allows one to assume a reflective attitude toward one’s socialised fate and to revise one’s self-understanding (FHN, 60).

Habermas suggests that the minimal self-understanding of ourselves as members of the species presupposes that we are all authors of our own lives and that we deal with others as authors of their own lives. By depriving the fusion of two sets of chromosomes of its contingency, intergenerational relations lose their naturalness which has so far been part of the taken-for-granted background of our self-understanding (FHN, 62). Without maintaining the distinction between the grown and the made, we risk a series of cumulative intergenerational violations of interactions. As Gadamer has shown, the formation of cultural traditions unfolds in the medium of questions and answers (Gadamer 2004). But genetic programs give future generations no opportunity to participate in this dialogue. Habermas doesn’t want this dialogue to be forced by economic-technological imperatives in a context in which the modern lifeworld simply lacks resources to articulate a response.

There are a number of objections to Habermas’ argument which suggest that questions around genetic interventions cannot be easily settled. In this vein, Bostrom comments that “had Mother Nature been a real parent, she would have been in jail for child abuse and murder” (Bostrom, 2007) Clearly there is a vast pool of facts which support the view that we can do better than nature, just as there is a vast pool of facts which support the contrary view. The set of facts that one draws upon perhaps depends on the prior unarticulated deep-seated pre-scientific convictions one holds. Such convictions may underlie Bostrom’s comment that “bioconservatives draw attention to the possibility that subtle human values could get eroded by technological advances …” (Bostrom 2007, 1). This trivialises Habermas’ argument. It is not “subtle human values” at stake but, according to Habermas, the set of fundamental distinctions that form the basis of autonomy, morality and sociality.

Habermas’ difficulty is in articulating these distinctions as criteria for acceptability of particular interventions. He distinguishes between acceptable and unacceptable interventions on the basis of “attitudes”, rather than what is actually done. Habermas can’t accept the instrumental attitude of parents fulfilling their egotistical dreams by designing their children, but can accept the defeasible anticipation of consent typical of an emergency doctor treating an unconscious patient (FHN 52).
Further, it seems that the consequences that Habermas predicts result from a genetically modified person’s belief (regardless of its truth) about the nature of intervention and their programmer’s attitude. One can surmise an instance of genetic intervention that remains unknown by the mature person, who consequently has none of these problems. It is therefore tempting to say we should simply find a way to adjust our responses. Mary Rorty suggests that if we find ourselves, as a result of genetic technology, at risk of regressing to a pre-liberal society where we discriminate between the real humans and not-quite real humans, we could then decide to maintain a society in which the bearer of genetic alterations had the rights of any other person (Rorty 2002). However, whether we can decide to change fundamental attitudes in this way is uncertain, especially when these fundamental attitudes are not those we currently have but can only be imagined in terms of our current attitudes. Perhaps we can neither speculate what our future attitudes will be nor prescribe what they should be.

In raising this issue, we see Habermas articulating a rational basis for vague concerns that are difficult to articulate, in order to bring them to consciousness in a way that enables public discussion. His sensitive interpretations of these everyday intuitions are directed at defending the lifeworld from an unconsidered and overhasty takeover by economic and technological imperatives. In this endeavour, he remains abstemious in avoiding strong ontological commitments. He is not concerned with any deep reality to do with the existence or non-existence of “fate”, “causality”, “contingency” or “autonomy”. Rather his concern is for “a form of life”, a grammar of inter-subjective relations conducive to the good life. He is concerned with what we do with language, how it ties in with our social practices not only whether, or how, it refers to objects recognised by natural science.

5 FAITH AND KNOWLEDGE

The idea of assuming virtual consent from the unborn is only one difficulty in resolving questions about genetic intervention. Given the scarcity of determinate facts or reasons which can serve as elements of a compelling argument, it is not surprising that Habermas turns to religious language to tease out his intuitions. His interest in religion is not new. In the 1970’s, he analysed religion primarily from a sociological perspective. By the end of 1980’s he acknowledged that religious images help to cope with the experience of contingency and provided comfort in times of need. (Habermas et al 2010, 15; Muller-Dohm, 384). More recently he notes a “growing political influence of religious orthodoxies” in the form of aggressive Islamism, Hindu nationalism, or Protestant fundamentalism (BNR, 1). The potential for conflict raises the issue of post-secular societies which, whilst largely secularised, must reckon with the continued relevance of different religious
traditions. Habermas doesn’t consider the existence of religious traditions within secular modern society as a zero-sum gain, a battle between the capitalistically unbridled productivity of science and technology on one hand, and conservative forces of religion on the other. Rather than one side gaining only at the expense of the other, Habermas invokes a third force, the civilising voice of a democratically shaped and enlightened common sense. Common sense is autonomous and can resist both naturalistic reduction and religious traditions, from which it can nevertheless learn.

Where religion is seen to conflict with science, Habermas argues that philosophy must take for granted the priority of scientific claims about the objective world. Scientific claims about the social world don’t necessarily have priority. He is critical of any instrumentally reduced notion of reason that is not interested in its own history. “This form of radical naturalism devalues all types of statements that cannot be traced back to empirical observations, statements of laws, or causal explanations, hence moral, legal, and evaluative statements no less than religious ones” (BRN, 141). Whilst post-metaphysical philosophy doesn’t dispute theological affirmations but asserts their lack of determinate meaning, Habermas doesn’t rule out the possibility that religious claims contain important universal moral truths. Avoiding sterile arguments such as the status of God’s existence, he looks to the pragmatic consequences of such commitments. “Philosophy, even in its postmetaphysical form, will be able neither to replace nor to repress religion as long as religious language is the bearer of semantic content that is inspiring and even indispensible, for this content eludes (for the time being?) the explanatory force of philosophical language and continues to resist translation into reasoning discourses” (PMT, 51). Habermas suggested an answer to the parenthetical question three years later. “Secular languages which only eliminate the substance once intended [by religious language] leave irritations” (FHN, 110).

As we have seen, scientific descriptions become highly disconcerting for commonsense as they come closer to our bodily and mental existence. Habermas suggests that “scientific theories change the content of our self-understanding.” They don’t “touch on the framework of our everyday knowledge, which is linked to the self-understanding of actors and speakers” - fundamental categories such as distinctions between persons and objects, what is and what should be, reasons and causes (FHN, 105). But what happens as we progressively subsume ourselves under scientific descriptions? To describe actions in terms of the generation of mental states by biological machinery misses what is relevant to what we recognise as agents. The scientistic belief that science will one day not only supplement, but replace the self-description of actors as persons by an objectivating self-description is not science, but

---

130 This represents a shift in Habermas view from TCA (in which he adopted the Weberian account of religion withering away in an enlightened world) to BNR, in which he accepts of the persistence of religious beliefs and practices (see also RR 79)
bad philosophy (FHN, 108). By its insistence on rational justification, science is certainly supported by a scientifically-informed common sense which has found its place in modernity. However, “no science will relieve common sense … of the task of forming a judgment, for instance, on how we should deal with pre-personal human life under descriptions of molecular biology that make genetic interventions possible” (FHN, 108).

What then is the place of religion in relation to scientifically-informed commonsense in a post-secular society? Habermas rejects the sharp liberal distinction between reason and religion. Non-believers should shoulder equal responsibility for undertaking “cooperative” translation of religious to secular language (BNR, 131). Whilst not a straight-forward task, translation can enable divergent traditions to find a common language other than that of the market or science. For example, in the controversy of how to deal with human embryos, believers might evoke Genesis: “So God created man in his own image ...” (FHN, 115). Here the notion of God’s creature expresses the equality of all humans. Habermas glosses this passage to suggest that God remains a “God of free men” only as long as we do not level out the absolute difference that exists between the creator and the creature. He interprets the concept of God as not needing to abide by the laws of nature like a technician, or the rules of a code like a biologist or computer scientist. Rather, God determines man by enabling and obliging him to be free. This contrasts to the entirely different causal dependence involved if the absolute difference assumed as inherent in the concept of God’s creation were replaced by a human who intervened, solely according to his own preferences.

Here philosophical interpretation brings religious language into “the universe of argumentative discourse … uncoupled from the event of revelation”. Translation doesn’t always succeed: “The metaphorical use of words such as ‘redemption’, ‘messianic light’, and ‘restoration of nature’ etc makes religious experience a mere citation. In these moments of its powerlessness, argumentative speech passes over beyond religion and science into literature, into a mode of presentation that is no longer measured by truth claims” (RR 75). And if it does succeed only by eliminating the substance intended, it leaves irritations. “The unbelieving sons and daughters of modernity seem to believe that they owe more to one another, and need more for themselves, than what is accessible to them, in translation, of religious tradition – as if the semantic potential of the latter was still not exhausted” (FHN, 111).

6 FREE WILL and DETERMINISM

We have seen that the distinction between the epistemic standpoints of participant and observer in Habermas’ analysis of the lifeworld and system and his mature theory of truth. This same distinction
is brought to bear on the longstanding debate about the possibility and nature of free will, a debate radicalised in recent decades by claims purportedly supported by neuroscience. Here we see Habermas reject reductionist interpretations of science, whilst leaving intact both the findings of that science and lifeworld intuitions of free-will. It is not surprising that Habermas defends the concept of free will. The critical theory tradition is distinguished by the idea that we are free to choose the good life, however circumscribed that freedom actually is. Habermas however, doesn’t neglect the emancipatory possibilities made available by the “external” perspective of a scientific observer. The importance of this perspective is also a crucial insight of critical theory. He notes that scientific progress has enabled an ongoing decentering which relativises the place of humankind in the world in enlightening ways – for example, the objectivating self-descriptions arising from Freudian or sociological reflection (Habermas 2007, 23-4).

Similarly neurological descriptions can help us recognise, and take account of, the ways in which the brain, as the physical substrate of rationality, can be causally impacted. By bringing this into view, objectivating neuroscientific accounts can enable the critical engagement with causal biographies and possibly, by the free will still presupposed, liberate people from certain internal constraints. Such external perspectives can also challenge views about responsibility for certain behaviours which we come to understand as consequences of neurological disorders, rather than moral failings. What is at stake for Habermas is that claims about free will touch on issues that have political dimensions requiring reasoned public deliberation. However, this deliberation is short-circuited by reductive naturalism which assimilates practical questions to considerations of causality which, it assumes, yield a “true” human nature.

Habermas declares his critical target as “the tendency to jump to philosophical conclusions from a successful and undisputed scientific enterprise” (Habermas 2007, 84) The tendency is seen in attempts to explain free-will as an illusion. He notes that we can understand the temptation to “explain” freewill in terms of three competing intuitions. As knowing subjects, our experience, backed by the systematic authority of the natural sciences, shows us that everything that happens is caused to happen according to regularities. As agents, we are convinced of the irreducible distinctiveness of the causal effectiveness of our minds – we can decide to intervene in the world in a way that is entirely up to us. As scientifically enlightened persons, we are convinced that the universe is one and includes us as part of nature and, in the light of this, the first two intuitions seem contradictory.

131 This view is strongly implied, if not argued for, in the work of neuroscientists such as V. S. Ramachandran, philosophers such as Paul Churchland and evolutionary biologists such as Steven Pinker (Ramachandran 2004, Churchland 1984, Pinker 2002).
This seeming contradiction has become acute in modernity, radicalised by ever more thorough objectifications of the human mind. This is evidenced in the call to replace the lifeworld’s “manifest image” with the “scientific image” (to use Sellars’ terms) of man. Against this call, Habermas suggests that two perspectives, that of both a participant in inter-subjective practices of reasoning and an observer of an external world of causes, are inextricably interwoven and run in parallel (Habermas 2007, 35-6). We have seen Habermas’ colonisation thesis evoke modernity’s tendency towards fragmentation and over-specialisation, undermining our grasp of the whole and alienating us from our critical aptitudes. Habermas also argues that reason bifurcates into two stances, whereby the objectifying stance characteristic of science takes precedence over the stance of the participant in the lifeworld.

Free-will is typically framed as a matter of an agent having a certain sort of causal efficacy, mental causation. However, what Habermas sees as essential to the concept of free will is not mental causation, but the internal link to reasons bound to social practices. The pawns in a chess game are, according to natural science, only bits of cellulose. But a chess player is not pretending that they are pawns. The concept “pawn” is bound into our practices. Similarly, reasons for action are real, although they only exist in relation to the social game of giving and asking for reasons. Free will, Habermas claims, is a natural part of the social world, contrasting to determinism not as mental causation, but as the ability to respond to reasons. As participants, we have reasons for our actions and tacitly assume we could do otherwise and it is up to us how to act. Free-will is a presupposition of the language game of responsible agency which reveals itself only to participants who take up performative attitudes vis-a-vis second persons (Habermas 2007, 15). Actors are aware that, as responsible agents, they are always already operating within a space of reasons towards which they should be responsive. The intuition of having free will is reflected in the presuppositions of this language game accompanying our actions. These presuppositions remain inaccessible for the observer, whose viewpoint is that of the uninvolved third person.

Habermas contrasts causal and rational explanations, arguing that reasons cannot operate in the manner of natural causes. The reasons which render actions explicable are different in kind from the events linked by laws of nature. To the extent that we are guided by reasons, we submit to intersubjectively shared norms (2007, 17). Unlike causes, reasons have a degree of indeterminacy. They are not decisive but can be either better or worse reasons. Reasons are reasons for us. It does not follow from rational explanations that any given person would reach the same decision in identical antecedent conditions. To be a responsible agent one must be motivated by letting one’s judgement be determined by reasons. As an “author” of one’s actions, one must take the initiative and attribute it

132 I borrow this example from Anderson (2005).
to oneself. Thus an actor is free when she acts on reasons she has made her own, but could also have acted against her better judgment (BNR, 159-60). In the process of deliberation, we must assume that the outcome of action is not pre-determined from the outset. This assumption is the very meaning of deliberation since without it, deliberation would be purely epiphenomenal, performing no function. Habermas stresses that since a will is formed, however imperceptibly, *in the course of deliberations*, we experience ourselves as free only in the actions that we perform more or less consciously (BNR, 155).\(^{133}\)

The open-endedness of a free decision doesn’t preclude it being *rationally* conditioned. The actor is free when she wills for reasons that, on deliberation, she accepts. We only experience as a lack of freedom, the external constraints that force us to act otherwise than we would on reflection (BNR, 157). Freedom involves an implicit endorsement of our conditioning. In the course of her deliberations, the actor reaches a rationally motivated position that is neither arbitrary nor a causal process but the result of rule-governed inferences. Habermas draws on the internal connection between reflection and freedom found in Kant’s sense of autonomy. But rather than adopting an unconditioned noumenal causation, cut off from all empirical contexts whilst intervening in the world, Habermas sees freedom as conditioned by being embedded in the context of reasons as they arise in the lifeworld (2007, 19). To be free is to bind ones will by being sensitive to culturally transmitted and socially institutionalised reasons. This can be contrasted to unambivalently unfree actions - compulsive, habitual, chance or neurotic actions which are typically described in terms of causes. I am unfree “if my decisions were determined as a neural event in which I was no longer involved as a person who takes a position” (BNR, 158). It is only by not noticing the slide from participant to observer perspective that it seems that the rational motivation of an action forms a bridge to determination of action by observable causes.

The contrast between reasons and causes is clearly seen by the fact that the causal explanations of natural science are appealed to in pleading that a person has acted unfreely and therefore cannot be held responsible. Habermas notes that in legal discourses, naturalistic explanations are appealed to if actions are unintelligible on the basis of comprehensible motives (Habermas 2007 pp. 19). As long as we are only talking about *limitations* to free will, that which is limited remains presupposed. But free will entirely disappears when action naturalistically bypasses the propositional attitudes of the actor, to be traced solely to nomologically determined events. Behaviour is then seen to be not decided by persons but rather fixed by their brains. From the neuroscientists’ perspective, “decisive arguments

---

\(^{133}\) We also act freely, when we remain responsible for actions about which we *should have reasoned* correctly, but failed to do so. Habits that bypass our deliberative scrutiny are still deemed to operate with our implicit agreement.
within the hierarchy of reasons can only ratify … what has already been long decided in regions of the brain far from consciousness” (Habermas 2007, 20). This form of explanation “requires switching perspectives from being the participant who accuses or justifies to being the observing analyst who … explains the behaviour” (Habermas 2007, 21). If the distinction between what were once complementary patterns of explanation is lost, the naturalistic explanation of behaviour loses any link to the norms of responsible agency (Habermas 2007, 22). Free will is eliminated and the language game of responsible agency collapses on the assumption that unconscious brain states completely determine all mental states.

John Searle objects to this account by arguing that the language games of neuroscience and of responsible agency are not in conflict, but are rather “different levels of description of the same system” (Habermas 2007, 71). The problem is that Searle’s claim simply begs the question of what this “same system” consists. The objects of neuroscience have colour, weight, texture, occupy space within the skull and are connected causally. The objects of mental states have no colour, weight or texture, don’t occupy space but refer to things in the world, are connected inferentially and are governed by social norms. It is only as a participant that we can pick out these states, after which the neuroscientist switches to the observer’s stance to note neural correlates. But why should this be an account of the same processes or “same system” Habermas responds that “we could be sure that we are dealing with descriptions of the same processes only if we could translate equivalent statements from the one language into the other …” However, “the language we employ for psychological processes and semantic matters cannot be reduced to physicalist or behaviourist language” (Habermas 2007, 89).

If we assume that the neuroscientific debunking of free will is correct, is it even possible that we adapt our normatively moulded consciousness to an objectivating self-description, “according to which one’s thoughts, intentions, and actions are not just instantiated by brain processes, but completely determined by them?” (Habermas 2007, 23). Whilst we may think that we are deliberating freely about reasons and that we could choose to act otherwise, can we persuade ourselves that our actions and thoughts are “really” caused by a series of electro-chemical reactions in our neurons in which we don’t recognise ourselves? Or is the participant’s perspective in the game of giving and asking for reasons unavoidable? Are there limits to self-objectification?

The conception of ourselves as persons depends on a distinction between doing and occurring. Neurological description abolishes this distinction by dropping any reference to normative success or failure in both action and reasoning, and replacing it with a language of events that simply occur (Habermas 2007, 26). Whilst an objectivating description of a one-sided determination of the mind by the brain might undermine an illusion, it also “dissolves this perspective from which alone an increase
in knowledge could be experienced as emancipation from constraints” (Habermas 2007, 24). By insisting on a strictly neurological account of the mind, researchers lack the resources to understand what it means to confirm or refute their own theories in the light of reasons. Such a stance amounts to a “performative contradiction”, much like someone trying to convince a sceptical opponent that the totality of human interactions, including their own deliberations, is determined in advance (Habermas 2007, 24). The limit to self-objectification is met when persons describe their actions as spatiotemporally identifiable events that can be explained nomologically. Persons can no longer recognise themselves as persons.

Habermas regards the explanatory models, terminology or language of the two perspectives of participant and observer as irreducible. “Descriptions of persons and their thoughts or practices cannot be translated into behaviourist or physical terms without losing or changing their meaning” (Habermas 2007, 25). This is glossed over in neuroscientific accounts, by the metaphorical assimilation of reasons to causes. Knowledge and its acquisition are irredeemably normative in a way that resists all attempts at empiricist redescription. The participant perspective can’t “be aligned with, and subordinated under, the observer perspective in such a way that we can capture ourselves, in an objectivating manner and observe ourselves from a fictitious view from nowhere - not just as acting and speaking subjects, but also as epistemic subjects engaged in the act of investigating ‘ourselves’?” (Habermas 2007, 26-7) The failure to recognise how deep-seated the participant perspective is can be traced to a basic scientistic assumption that the objectivating perspective of the natural sciences has priority over the participant’s perspective.

Habermas sees neither epistemic perspective as having priority, since “the perspective of an observer who adopts an objectivating stance towards something in the world, is a fortiori interwoven with the perspective of participants in discourse who, in presenting arguments, adopt a performative stance toward their critics” (BRN, 169). This means that the observer’s objective reality can only be constituted together with the inter-subjectivity of possible communication. Our participant’s perspective is not something we can step out of, to see from the outside. The resistance to the naturalistic redescription of our self-understanding as persons doesn’t reflect a mere illusion, but is explained by the fact that “there is no getting around a dualism of epistemic perspectives that must interlock in order to make it possible for the mind, situated as it is within the world, to get an orientating overview of its own situation” (Habermas 2007, 35). “Objective” descriptions do not issue from isolated minds reflecting the world “as it is”, but constitute claims on others, even if those others are an imagined ideal audience.
For Habermas, both perspectives arise from adaptations as our species acquired communicative competency. On this account, archaic humans included the objective world within the social relations of the inter-subjectively shared lifeworld. As they sought to master unpredictable nature, humans developed a language of objects and events that enabled increasingly de-socialised accounts of objective nature (Habermas 2007, 35). Two distinct patterns of explanation developed, one reflecting the participant perspective, the other the observer perspective. Scientific disciplines developed reflecting this bifurcation, with the natural sciences adopting “explanation” and the social sciences and humanities adopting “understanding” as their respective methods (Habermas 2007, 36).

In the 17th century, philosophy was confronted with the question of “what it means for humankind to understand itself in the context of scientifically objectivated nature” (Habermas 2007, 36). Here, in what Foucault called “the analytic of finitude”, we see the objectivating stance of the observer looping around to focus on human nature – subjectivity itself – as an object of science. With the development of sciences such as neurology leading the call to replace the manifest image with the scientific image of man, the interwoven perspectives which once ran in parallel came to be seen as conflicting. Whilst recognising the emancipatory possibilities of the observer perspective, Habermas rejects the scientistic claim of its priority. He wants a more liberal form of naturalism that is not restricted to the objects of natural science but embraces, for example, the mind, beliefs, reasons, goals, meanings and morals.

In our everyday practices there is an allocation of labour, whereby naturalistic objectivating explanations are employed for actions that cannot be made sense of from the participant’s perspective. It is only by the extension of the logic of natural science to “explain” action and reasoning per se that a metaphysical leap is made. In claiming that reality is restricted to the scientific image, the reductive naturalist forgets that the objective perspective is itself interwoven within the normative space of reasons, inaccessible to natural science. The natural sciences cannot embrace all we encounter.

7 SCIENCE AND PHILOSOPHY

We have seen how Habermas’ “weakly transcendental” project adopts an empirical starting point – the rational reconstruction of fundamental presuppositions implicit in communication. Philosophy becomes interdisciplinary, drawing on the sciences, whilst retaining a distinct role by its connection to the lifeworld. Habermas’ critical philosophy reveals distortions of the lifeworld by providing scientific perspectives from outside the everyday lifeworld. Although a provocation to the lifeworld’s...

---

134 Habermas recapitulates a line of reasoning from Sellars (1963)
conservatism and opacity, science can’t claim automatic authority but must ultimately be accommodated within the lifeworld’s commonsense. It is philosophy that has the capacity to broker this accommodation. At the same time, science can demand that philosophy make its positions perspicuous, which it must do in the light of both science, which must be taken seriously and common sense intuitions, which can’t be abandoned. The relationship of philosophy to science is two-way. Whilst science challenges the lifeworld, philosophy draws on the lifeworld to challenge science’s occlusion of perspectives and validity spheres beyond its limits. Certain problems “are no longer visible from the perspective of … institutionalised science and its discourses” (TJ, 289).

Habermas would recognise science’s emancipatory potentials, for example, the treatment of genetic disorders, or neuroscience revealing the biological basis of behaviours we could otherwise only moralise about. What he criticises is the slide from science to bad philosophy seen in liberal eugenics, or reductive neuroscience, propelled by the momentum of technology and capitalism. Like the denial of free will, such innovations can run roughshod over the lifeworld’s intuitions of freedom and responsibility. Habermas draws on the semantic potential of religion to give voice to these lifeworld intuitions, at the same time, eschewing simplistic reconciliations. His sensitive interpretations of religious language and everyday intuitions are directed at defending the lifeworld from an overhasty takeover by systems imperatives. His concern is less about the objects recognised by the natural sciences and more about the pragmatics of communication and forms of intersubjectivity conducive to the good life.

These concerns can be traced back to Habermas’s fundamental commitment to democracy and the public sphere which he believes increasingly lacks cohesive perspectives. In the face of its fragmentation and colonisation, he mobilises the lifeworld’s resources to bolster the public sphere against the impoverishment of its communicative resources at the hands of scientism, the economy and bureaucracies. A question I will raise in the final chapter will be the extent to which Habermas’ theoretical structures restrict and limit his analyses. Habermas’ reconstructive sciences don’t venture far from the idealised realm of communication, to lay out as Foucault does, a genealogy of actual things and bodies. We will see both Habermas and Foucault restricted by stances on which their analyses nevertheless depend.
Chapter 8
Science Modernity and Freedom

Foucault and Habermas both see science, particularly the human sciences, as in need of critique. Their concern is not science’s truth or falsity, but its potential to distort other discourses and practices in a way that limits freedom and facilitates domination. Rather than a misapplication of science which in itself is “pure” and disinterested, they are concerned about a misevaluation which boosts science’s authority by proclaiming its purity and disinterestedness, so insulating it from challenge by aspects of validity it suppresses. But as we will see, there are differences. Following the so-called “debate” between Foucault and Habermas, commentators tended to support or reject their respective positions. More recently, others have sought to reconcile their best insights (Allen, 2008; Cooke, 2006; Koopman, 2013). I don’t want to either adjudicate or reconcile, but simply account for the tension between their two projects. This concluding chapter will tease out the conjecture of a broader framework, which reveals the positions of Foucault and Habermas as two distinct tendencies within the structure of modern thought. Their commitments are revealed not so much by what they say directly about science, but how they use it, the role they grant science in their own critique of reason.

I shall start by considering self-referentiality and its radicalisation within modernity in order to initially characterise what I call strategies of “discovery” and “self-transformation”, employed respectively by the two thinkers. I will examine the imbricated dialectical relationship between these strategies by considering them in terms of participant and observer perspectives and then necessity and contingency. Having constructed this framework, I shall examine some confusions and controversies surrounding the role of normative foundations in Foucault’s and Habermas’ work, and the tensions between power, freedom and rationality. I will extend this analysis to bear on the salient methodological features of Foucault’s genealogy and Habermas’ evolutionary theories and their differing accounts of the development of science.

1 SELF-REFERENTALITY AND ITS MODERN RADICALISATION

Both Foucault and Habermas recognise the conditioning of their thought by historical, cultural and social factors. They see modernity characterised by self-referentiality, and craft their critical projects in light of this. Their problem is how to practice philosophical critique given this self-referentiality. By “self-referentiality” I mean the tendency to take one’s thought or practices as an object of thought. In so doing, one suspends commitments from beliefs, treating them as hypothetical, in order to examine them. Habermas thinks that “modernity … has to create its normativity out of itself” a task
inaugurated by what Foucault identifies as “the permanent reactivation of an attitude … a permanent critique of our historical era” (PDM, 7; PT, 109).

Whilst humans have presumably always been able to adopt a third-person stance on themselves, both Foucault and Habermas see this as radicalised in modernity. Foucault notes that the modern cogito can no longer infer from “I think” to “I am” because it is not clear whether “I” is what I consciously survey, or what is unthought that conditions, such a survey (OT, 352-4). This instability produces a craving for explanation that generates a never-ending compulsion to keep uncovering what remains hidden in our nature. Foucault’s analysis is directed as much to the sciences as philosophical reflection. Whilst philosophy motivates a desire for thinking the unthought, it is science that provides the content. In the contemporary scene, specialised fields such as neuro-physiology and evolutionary psychology most explicitly uncover the conditions of possibility for our thinking, whilst leaving more unthought which conditions this uncovering. Whilst never fully grasped, “the whole of modern thought is imbued with the necessity of thinking the unthought” (OT, 356). This necessity to pursue and uncover a background of continually receding conditions of possibility is “profoundly bound up with our modernity” (OT, 357).

We moderns assume the possibility of reflectively questioning received beliefs and values, and gaining critical distance from inherited norms and roles. In Chapter 6, we saw Habermas trace this tendency in his social evolutionary account as potentially emancipatory differentiations accruing within the lifeworld. Yet the heightened contrast between participant and observer perspectives can be problematic. We saw in Chapter 7, Habermas’ contrast between the stance of a participant in social practices and the third-personal objectivating stance of a neuroscientist towards that participant. He is concerned that fundamental categories are increasingly threatened by interpretations of science which proclaim the primacy of the scientific observers’ perspective (Habermas 2007, 36). Within modernity, these two moments, the participant and the observer, have been radicalized to be seen as two conflicting perspectives. Whilst able to debunk another’s arguments by casting aspersions on their reasoning capacity we still, to some extent, remained within the “space of reasons”. But now the stakes are raised. By adopting the scientific observer’s stance to employ the authoritative language of causes, we deny epistemic normativity to not only the irrational thinker, but seemingly, by extension of the consequences of the same argument, to ourselves.

Similar difficulties had been recognised by the previous generation of critical theorists, who in highlighting the dominance of instrumental reason, left little room for critique to stand. Is there any position from which we can confidently commit to our judgements? Can we find an “external” perspective sufficient to critique conventional reason? These are the questions that Habermas and Foucault inherited in responding to modernity’s self-referential turn. They both criticise the stance of
the detached scientific observer who proclaims its primacy. Yet both, in different ways, must draw upon this stance. My conjecture is that their responses, whilst addressing a bifurcation of reason characterised by participant and observer perspectives, unavoidably reflects this structure as two “moments” of inquiry which I will elucidate as a “strategy of discovery” and a “strategy of self-transformation”. These responses arise from the life experiences that have animated and ultimately shaped the projects of Foucault and Habermas, to which I will turn to gain an initial foothold on their respective strategies.

2 DISCOVERY AND SELF-TRANSFORMATION

In Chapter 1, I mentioned that Habermas has lived through the tumultuous years of Germany’s defeat, the allied occupation and war crimes trials, the drafting of a new constitution and Germany’s social, economic and industrial reconstruction - much of the time against the backdrop of the Cold-War. He was shocked not only to learn of Nazi atrocities, but to see some of his elders criticising the trials of perpetrators (Habermas 1983, 57). It is easy to understand that Habermas would be anxious to defend a form of reason capable of defending itself against the forces of irrationality which had overtaken his country. Unsurprisingly, he invested cautious faith in a project of extending and clarifying reason and science in an attempt to anticipate their limits and possibilities. Scientific technology promised freedom from material wants, and human and social sciences better understandings of the crises faced by modern societies. Yet he remained aware of the risk of new forms of enslavement, brought on by the narrow instrumental reason his Frankfurt School predecessors alerted us to. He turned to democracy which he analysed in terms of a theory of communicative rationality, in order to promote and justify its more adequate conception, based on an ethics of discourse. Whilst taking science seriously, Habermas’ prior commitment is to the lifeworld and its communicative structures that he sees being eroded by scientism, technocratic consciousness and systemic imperatives. He wants to safeguard the lifeworld’s structural mechanisms so vital to open and free societies. In this endeavour, he will pursue a strategy of discovery of what is stable and universal and, in a qualified sense, necessary.

Foucault’s biography is equally emblematic of the 20th century. As a gay man growing up in 1940’s France, we can imagine the internal conflicts generated by coming to understand and articulate his sense of self, as opposed to accounts of scientific “experts”. In response, his work sought to liberate the human capacity for self-formation and choice. His genealogies portray knowledge and power inextricably locked together to produce categories, objects and subjects. Despite its omnipresence, power/knowledge is contingent and always susceptible to resistance. In this spirit Foucault saw his challenge as revealing “what is given as universal, necessary, obligatory, what place is occupied by
what is singular, contingent, and the product of arbitrary constraints” (PT, 113). Having lived through the unrest of 1968 he was, like Habermas, well aware that rebellion may not lead to emancipation, but just new forms of domination. So his challenge is not to say that “everything is bad, but that everything is dangerous … the ethico-political choice we have to make every day is to determine which is the main danger” (EW1, 256). Foucault is pointing to the constant risk of the authoritative truths of science being drawn into constellations of power/knowledge in which they have negative consequences. What is dangerous may or may not turn out to be bad, but we had better keep alert to the risk. Rather than seeking the ideal in order to eliminate all danger for all time, we should operate from our provincial perspectives to discern the main danger every day. This discernment occurs in specific contexts and cannot be further legitimated by any appeal to science or general principles in abstraction from that context (PT, 115).

Foucault did not, like Habermas, articulate theoretical structures in order to posit regulative ideals for the good life. He didn’t have that much faith in reason. Faith itself was a form of complacency, against which we must always be vigilant. Wherever knowledge is too secure, it must be challenged by revealing its origins – who constructed it, for what purposes and with what interests. Foucault could be said to follow a strategy of self-transformation directed towards “no longer being, doing or thinking what we are, do or think” (PT, 114). Whatever is given as natural – madness, delinquency, sex, even truth and morality - is the way it is only because our conceptual and practical activities have constructed it that way. What we thought was necessary is revealed as contingent and this revelation provides freedom for new orders to emerge. Rather than affirming change for change’s sake, Foucault is urging an ongoing struggle of spiritual transformation by problematising our most secure commitments to open up fresh perspectives and freedoms.

For Habermas, the possibility of freedom is provided by certain unavoidable aspects of human existence such as the norms of communicative action, required to make sense of the world. Embracing a sliver of certainty, his strategy of discovery seeks what is stable and universal. Foucault challenges what we think is stable and universal. By revealing what we thought necessary to be contingent, Foucault’s strategy of self-transformation loosens states of domination (EW1, 283). Foucault not only suspends commitment to truth, but the conceptual framework of ahistorical universal categories which precede truth. 135 Whilst both Foucault and Habermas are broadly within the Kantian Enlightenment tradition, they engage in two contrasting forms of philosophical reflection on the limits of thought and action.

135 See Rabinow and Dreyfus 1983, 49-52 for a discussion of Foucault’s suspension of truth and meaning.
Habermas’ approach is derived from Kant’s concept of *critique* in its recognition of the limits of reason and its critical-transcendental power to ground claims to truth and normative rightness. Substituting a fallibilistically conceived philosophy of inter-subjectivity for Kant’s philosophy of the subject, Habermas wants to rehabilitate the essential kernel of Enlightenment humanist thought, in which he sees a liberating communicative openness. Modernity is the incomplete project which must not be abandoned (Habermas 1996, 38-55). He identifies himself with the Enlightenment project, “consist[ing] of the relentless development of objectivating sciences, of the universalistic foundations of morality and law, and of autonomous art, all in accord with their own immanent logic”, a process which “encourage[s] the rational organisation of social relations.” (Habermas 1996, 45) Committed to these fundamental Enlightenment values, Habermas wants to surpass the negative critique of the previous Frankfurt generation. He incorporates the sciences into a universalistic theory in a way that contributes to the protection of democracy and rationality from threats that lie within modernity. We have seen how Habermas reconstructs the rational internal structure of communication to reveal modern rationality’s capacity to distinguish three validity claims corresponding to three worlds, along with their associated attitudes and value spheres. Participants in argumentation cannot avoid certain ‘idealising presuppositions’. Whilst the process of enlightenment of modern societies involves the lawful use of reason to reconcile the ideal and the real, this potential is obstructed by one-sided rationalisation. Referring to Foucault, he identifies a further threat as the “radical critique of reason” (PDM, 336-7).

Foucault identifies with Kant’s Enlightenment *attitude*, as exemplifying a certain form of reflection on the present. He sees this form of reflection within a stream of Enlightenment critique in which reason perpetually questions the present and hence itself. We must not submit to the “blackmail” of the Enlightenment, the demand to be either “for” or “against” it (PT 109-10). Embracing the self-referentiality of modern reason, Foucault rejects the notion that “one either recognises reason or casts it into irrationalism – as if it were not possible to write a rational criticism of rationality” (Foucault 1989a, 353). Since the Enlightenment is part of “the historical ontology of ourselves” that has determined who we are, any position for or against it remains inherently partial and conditioned by this determination. Enlightenment requires a philosophical ethos of permanent critique (PT, 109). The critical question today is “how can the growth of capabilities be disconnected from the intensification of power relations?” (PT, 116). This requires a process of becoming otherwise than we are through the *agonic* use of reason, to call into question what is given as reason. By neither endorsing nor rejecting, but *suspending* the normative framework, Foucault adopts the stance of the radical observer who inquires into reason, not in terms of its nature, limits or foundations, but in terms of the contingency “its history and its geography … its immediate past and its present reality” (Foucault in Canguilhem 1991, 9). Foucault’s histories show that what appears as the one and only possible reason has a history. He doesn’t think we need a theory of what rationality *really* is, in contrast to what it
takes itself to be. Social practices exist within certain “regimes of rationality”, certain historical forms which cannot be seen in relation to a perspective outside all regimes.

Foucault sees humanism as an ambivalent force that excludes as much as it includes and constrains as much as it enables. However, by criticising humanism, Foucault isn’t challenging its principles so much as its pretensions to universality and ahistoricity. “What I am afraid of about humanism is that it presents a certain form of ethics as a universal model for any kind of freedom. I think there are more secret, more possible freedoms, and more inventions in our future than we can imagine in humanism” (Foucault 1988, 15). Foucault would argue that by making universals that have emerged in history less vulnerable to historicity, humanism stultifies the “undefined work of freedom”, by which we transform ourselves. Far more wary of reason’s ahistorical universals than Habermas, he doesn’t offer a theory of what rationality is, but challenges particular forms of rationality, in a way that his critics have seen as undermining reason per se. Foucault’s genealogy articulates a conception of critique distinct from Habermas. As Owen argues, Habermas’ critique legislates an orientation to a transcendental ideal in terms of a reconciliation of the real with the ideal by the lawful use of reason. Foucault’s genealogy exemplifies an orientation towards an immanent ideal in terms of becoming otherwise than we are through the agonal use of reason (Ashenden & Owen, 21-45).

Both Habermas and Foucault are aware of the tension between their strategies, which they see in some sense as problematic or paradoxical. Foucault wants to transform us by his critical histories that exhibit the singularity, contingency and arbitrary constraints of our forms of subjectivity, so enabling us to think and act differently. Habermas discovers a universal form of subject, a modern decentred subject, which he employs as a regulative idea. Foucault wants to show that what Habermas puts forward as universal are really singular and contingent, products of arbitrary constraints. Habermas wants to show that Foucault is caught in “a self-referential denial of universal validity claims” (PDM, 286).

By his strategy of discovery, Habermas is a participant in a game of truth that seeks what we are, what society is, as something determinate, stable and universal, pre-existing discovery. With his fallibilism distinguishing his stance from transcendentalism, Habermas’ reconstruction of developmental schemas reveals this way we are, which future empirical science will either confirm, or show that we are some other determinate way. Foucault’s strategy of self-transformation adopts the stance of a distant observer who has no commitment to our norms or conceptual categories. What appears as the most fundamental categories of nature is our own creation. Species like the homosexual, the delinquent or the insane are, in all their diverse historical incarnations, our inventions. Our norms, conceptualisations, naming and categorising, interact with the things normed, conceived, named and categorised, and not only things, but people, classes, kinds of people and ideas.
Foucault is concerned with how subjects are constituted. Just as there is no thing-in-itself of pure madness, there is no pure subject, no “I” or “me”, prior to the descriptions given to a person. Nothing, including our self-descriptions, is either one thing or another except that history made it so. And this is a history of struggle. “The history which bears and determines us has the form of a war rather than that of a language: relations of power, not relations of meaning” (PK, 114). Every way in which I can think of myself as a kind of person has been constituted within a web of historical events. Foucault wants to “discover how it is that subjects are gradually, progressively, really and materially constituted through a multiplicity of organisms, forces, energies, materials, desires, thoughts etc.” (PK, 97). This multiplicity reflects the three axes of Foucault’s analysis - knowledge, normativity and self-relation. Our experiences are “fabricated” in the sense that they are put together from historically sedimented systems of knowledge, sets of rules and norms, and ways of self-relating. (EW 3 243)

Both Habermas and Foucault were aware of these two moments, participant and observer, under various guises invariably picked up by commentators – universalism/nominalism, mind-independence/construction, necessity/contingency, foundations/practices, objectivity/consensus, transcendental/empirical, and so forth. These pairs don’t all reduce to the same underlying phenomena. However, the projects of Foucault and Habermas can be characterised in terms of two seemingly irreconcilable moments suggested by these terms which line up more or less consistently with one or the other strategy. I say “more or less” because these stances are unavoidably imbricated within each other, one explicit and thematised, the other implicit and suppressed.

3 PARTICIPANT AND OBSERVER

To see the two different projects in relation to each other, we need adopt a broader frame of reference characterised by the participant and observer perspectives of modern self-referentiality. By “participant”, I mean a participant in a set of social practices governed by norms. By “observer” I mean one who observes such practices, whilst suspending commitment to their governing norms. For example, I may be a participant in the social practice of morality, subjecting others, as others subject me, to praise and blame for actions for which we hold each other and ourselves responsible. An observer (for example a social scientist) “steps back” to view such practices whilst bracketing the governing norms. This scientist observer may note, for example, the social function of these practices. A similar participant/observer relation applies to the social practice of reasoning, where a neuroscientist or psychiatrist for example, observes the norm-governed activity of the reasoning subject, whilst bracketing that subject’s rational norms.
The perspectives are imbricated. The observer of social practices is never merely an observer but is, at the same time, a participant. Even the most “pure” scientific observer is a participant in the norm-governed social practice of science. By observing a particular practice as an object of inquiry, the scientist as participant takes for granted an immeasurably vaster expanse of unproblematic background practices. Whilst we can’t think of an observer’s distanced perspective floating free from a participant’s commitment, neither can we exclusively countenance a participant perspective. However, what it is to be a rational agent viewed from the participants’ perspective seems in tension with the causal conditions of such agency viewed from the observers’ perspective. This tension cannot be resolved in favour of either perspective as more fundamental. We have no alternative to shifting between the two perspectives. The perspectives of participants within a shared lifeworld and observers, who adopt an objectivating attitude towards the lifeworld’s objects, are inextricably interlocked (Habermas 2007, 35; OPC 376-7; BRN, 169). In terms of this framework of imbricated perspectives, Foucault and Habermas necessarily occupy positions of both participant and observer. The difference is the perspective that each will thematise whilst suppressing the other, and their attitudes towards the movement between imbricated perspectives. Foucault wants to maintain this movement by the ongoing process of a “rational criticism of rationality”, problematising whatever appears as ahistorical and universal. Habermas thinks we need to nail at least some things down.

Habermas thematises the participant perspective by explicitly declaring certain “foundations” on which his critique rests and which he urges us to accept. By “foundation”, I am not implying an epistemological theory which commits one to “foundationalism”, but simply a basis on which we should all agree, a point where argument stops so we can proceed with theory-building. Since everything is arguable we need certain points on which we can agree. For example, whilst we can argue over the ideal speech situation, the value of undistorted communication etc, these regulative concepts are put forward by Habermas as foundational in his theory. As a participant in the philosophical game of argument, Habermas uses these foundations to build a comprehensive social theory.

However, Habermas must firstly gain a view external to that of the everyday lifeworld. He therefore adopts an observer’s stance by incorporating science into his theory. By drawing on generalised empirical knowledge to reduce the context-dependency of understanding, he can analyse symbolically prestructured objects and events. His Frankfurt School predecessors failed to see that a theoretical account of social evolution can entail progress, seen in terms of learning. Rather than seeking progressive structures only in history, Habermas finds them in the pre-historical development of communicative capacities. The learning processes by which society evolves from pre-conventional, through conventional, to post-conventional stages is reconstructed in speculative developmental theories of human communicative competences. The discovery of universal features of
communicative action necessary to being the sort of creatures we are, restores critical theory’s critical ability. Sciences such as developmental psychology and anthropology provide plausible, though fallible, empirical universalistic theories which philosophy stiches together to bear on contemporary social crises. The developmental aspects of these theories bring into view the phenomena of social evolution - the “linguistification of the sacred”, the emergence of three discrete “worlds” each with their corresponding validity spheres, and the development of systems that co-ordinate action “behind our backs”. Combined with idealisations gleaned from speech act theory, this progression of developmental milestones provides Habermas’ normative standpoint from which he can diagnose the ills of modernity, such as the “colonisation of the lifeworld” or the “fragmentation of culture”. Such diagnoses would not be possible if not for this external perspective that science affords.

From his genealogical observer’s perspective, Foucault charts the course of concepts and discourses under different truth and power regimes as they interact with each other and the non-conceptual and non-discursive. His fine-grained nominalism reveals the plurality and contingency of events and things. Foucault generalises both knowledge and power. He goes beyond conventional boundaries of the sciences to undertake a “history of concepts” by analysing “discursive formations.” Power imposes conceptual structures on the world, leading to knowledge of it ordered by these structures. His generalised notion of all-pervasive power/knowledge enables us to see an ensemble of forces that relay power. Foucault is particularly concerned with the unassailable power-effects of the human sciences, which ripple through the social order, restructuring its relations to reconstitute both subjects and objects. By explicitly adopting an observer’s stance, abstracting from the question of truth to deal only with what is taken-as-true, eschewing commitment to scientific categories, Foucault can reveal the fine-grain of events beyond everyday consciousness.

As we will see, Habermas is concerned by what he sees as Foucault’s denial of epistemic and ethical commitments. Habermas’ own commitment is anchored in the idea of the decentred subject of modernity, a universal toward which the “logic of development” or learning process leads. Foucault however, was critical of philosophies which rely on a universal subject beyond question, especially products of processes of European modernisation that are “so universalising, so dominating with respect to others” (PT 115-6). He is “not prepared to identify reason with the totality of rational forms which have come to dominate” (PPC, 35). Today philosophy’s task is “to know how and to what extent it might be possible to think differently, instead of legitimating what is already known” (UP, 9). By starting from, rather than questioning, what is “already known” (the decentred subject), Habermas’ critique can’t criticise its own standard and thus stalls the movement between the imbricated perspectives of modern self-referential reason.
Habermas would respond by claiming that the differentiated rationality of the decentred subject can “burst asunder” the “provinciality” of other forms of subject (PDM, 322). Habermas might argue that this is obvious though, given fallibilism, not certain in any metaphysical sense. But if we can’t accept anything as sufficiently certain, we can’t claim anything. Against this, Foucault would insist that it is precisely what is taken to be obvious and certain that must be problematised. This has no finality. Modernity requires “the permanent reactivation of an attitude … a permanent critique of our historical era” (PT, 109). Foucault wants to show the decentred subject not as a regulative idea, but as one form of reason amongst other possible forms.

By adopting the observer’s perspective, Habermas uses science to reach determinate conclusions not available to the everyday lifeworld, and Foucault undermines assumptions which support scientised lifeworld conclusions. Like Habermas as user of science, Foucault as genealogical observer questions what was previously unquestionable. But Habermas ends up as a participant in a game of truth that seeks to discover what must be necessarily presupposed, in order to yield conclusions, however fallible and provisional, which have determinate practical application to the good life. Foucault sticks doggedly to his radical observer’s perspective to question what seems to be necessarily presupposed, by relativising such presuppositions as being presupposed only by creatures such as Man who, in any case, may soon disappear (OT, 421-2).

4 NECESSITY AND CONTINGENCY

The different strategies can be formulated in a way that heightens what is at stake – necessity versus contingency. Foucault wants to draw attention to the constitutive role of concepts, which open up the world in certain ways, rather than other ways. The world is radically contingent because we “invent” it, and could invent it differently. Habermas, more focussed on discovering what exists, would resist this. Certain standards, such as the pragmatic presuppositions of communicative action are necessary as standards of reason and argument (OPC, 21-105). This is conditional necessity. If we are to be the sort of creatures that we are – reasoning, arguing, progressing towards post-conventionalism - then we must accept certain presuppositions underlying our practices. Habermas is unequivocal that reconstructive sciences aim to elucidate underlying universal presuppositions as a condition of possibility of any communicative interaction whatsoever. Societies pass through a pattern of socio-cultural change, comprising a “logic of development”, a sequence in which no stage can be passed over and each stage presupposes previous stages (McCarthy 1978, 247). According to Habermas, the communicative competences and distinctions that emerge in modernity represent the determinate way we have become, or are becoming, according to the “logic of development” or learning process. This represents progress. The extent to which we depart from this logic which culminates in modernity’s
decentred subjectivity, is due to the “dynamics” of development, the empirical, contingent factors. With this expression “logic of development” suggesting progress as well as a form of necessity, Habermas aims at discovering both the truth of what we are, and what we ought to be. Whether this truth ends up involving Habermas’ idealisations and milestones, reconstructed from developmental psychology and anthropology, or something as yet completely unthought, there is still a way we are, and are becoming, which it is the task of science and philosophy to discover, although their claims will always be fallible.

Whilst Habermas’ notion of developmental dynamics recognises the way we are is not entirely conditioned, Foucault highlights this as instability. He would most likely suggest that Habermas’ principles and idealisations are themselves a further play of power and cannot have the status Habermas grants them. He is not interested in the determinate way the world and ourselves are or must be, which fallible science slowly uncovers. He is interested in the way in which our concepts order the world or “invent” it. Whilst we cannot simply choose the world we want, in some sense we contingently constitute the world and ourselves, as subject-objects within it. By showing how subjects-objects are put together, so that we could undo them, if we choose, Foucault reminds us of our possibilities and responsibilities.

My view is that what we call “necessary” or “contingent” is applied pragmatically in relation to logic, laws of nature, technological possibilities, practical obstacles, morality, convention, law etc to demarcate between that which we can and can’t (or don’t want to) know, control, predict or accept responsibility for. The difference between necessity and contingency is not a difference of kind but of degree of rational certainty best located in our commitments which, beyond a certain threshold, we call “necessary”. Habermas’ idea of a logic of development is not an empirical claim but an appeal to accept certain fundamental aspects of our condition as representing the direction of social, cognitive and rational progress. For example, in the course of their development, humans make increasingly finer differentiations between types of validity. By implying necessity, Habermas urges us to stop wasting time on useless possibilities and accept what appears reasonable as a stable foundation on which to build useful theories. If future science shows that this appearance is not the case, then we would need to accommodate whatever science shows is the case. Foucault’s discontinuous histories disrupt this picture. We can’t discern any logic of development, but only constant struggle. Progress or regress doesn’t even come into view for the distanced genealogical observer. Habermas’ trajectory of progress is only a projection of the human, all too human. And why should we endorse the sort of creatures we are? Foucault wants to open up possibilities. Whilst Habermas’ fallibilism admits that the world can always be some determinate way that our descriptions fail to capture, Foucault presses that the world is not any way at all, apart from our contingent descriptions of it.
Fallibilism suggests that the world is a particular way independently of our descriptions of it, with fallibilistic science approaching this way by falsifying descriptions that fall short of it. Foucault would point out that this claim is already a description of a particular way things are, not a direct experience of the noumenal bypassing description, but a particular description of the world imagined as independent and determinate. Whilst Habermas’ fallibilism readily admits that the world could be different than the way we think it is, Foucault might suggest that the world could be different if we think it is. Foucault suspends commitment to the “truth regime” – the framework of categories which precedes the division of the true from the false. In PP, he reminds us of a truth that isn’t already there, waiting to be discovered (PP, 235). Against the “philosophical-scientific standpoint”, Foucault recalls this archaic notion of truth that does not appear for all, but requires the subject to “modify itself, transform itself, displace itself, it must become … other than itself in order to have the right to access to truth.” (HS, 17) 136

Aiming to discover universal presuppositions for communicative action, Habermas’ fallibilism leaves room for correction by empirical science. However, by suspending commitment to any particular truth regime, Foucault’s strategy of self-transformation suggests not so much discovering what was hidden, but having one’s eyes opened to see new significances and connections. “Where the role of science is to make known that which we do not see, the role of philosophy is to make seen that which we already see” (cited in Kelly 2009, 129). He wants to reveal our constitutive activity as one particular form of subjectivity within constellations of power/knowledge. What is taken as truth changes not only in response to scientific discovery, but also in response to shifts in subjectivity. It is because the conceptual framework realigns terms to refer them to differently constituted objects, that Foucault can claim that only from the 19th century, “the homosexual [became] a species” (HS1, 43).

If I am correct that the difference between what we call “necessary” and what we call “contingent”, involves a pragmatic choice that depends on our purposes, it appears that the difference between the moments of discovery and self-transformation is also not a matter of truth or rightness. As a pragmatic question, there is nothing that can settle this apart from the purposes to which Habermas and Foucault put their inquiries. Clearly we must reach determinate conclusions on which we have confidence to act. And clearly we must continue to problematise. As we have seen, each analytical strategy requires a particular theoretical framework to get at its target. And each strategy brings certain objects into view whilst occluding others. We can employ this broader pragmatic frame to unravel some confusions and expose some limitations of both strategies which can ultimately be revealed as moments of modern reason which we struggle to articulate from our positions within it.

---

136 See Chapter 4.
In chapter 3, I raised one of the principal charges brought against Foucault - he fails to provide normative foundations. This much discussed criticism bears directly on the role of science. We have seen how Habermas employs science to discover principles necessary as universal presuppositions for communication. In contrast, Foucault’s histories seem judgemental, yet provide no secure foundations for their judgements. According to Habermas, Foucault’s judgements are “crypto-normative”, resting on normative assumptions to which he is not entitled (PDM, 276). It seems that Foucault’s call to “create ourselves as a work of art” ignores the intersubjectivity by which we are accountable (EW1, 262). Without any principled position, Foucault’s work itself can only be arbitrary.

For Foucault, norms are not laid out in advance, to guide the conduct of a future community. Rather, normativity is generated in the context of, and for the sake of, transforming and reconstructing existing problems. The very idea of a normative basis laid out in advance, is the object of Foucault’s critique. Like Wittgenstein, Foucault sees ethics as not a matter of principles, but interaction. By decentering the subject, Foucault similarly problematizes the idea of a stable, ethical self, guided by established principles. Falzon is correct to suggest that Foucault’s thinking is amenable to a dialogical interpretation (Falzon, 2010). The constitution of norms requires the constraints of an actual field of struggle and counter-struggle (STP, 3). Subjects are immersed in inherently reciprocal intersubjective relations. Foucault reveals justificatory frameworks emerging out of a history of dialogue that has no underlying logic. By not articulating foundations we don’t necessarily lose normative standards to which to appeal. We still follow rules, though not like tram tracks laid out in advance. Responding to specific contexts, there is no single way in which ethical agents arrive at a proper course of action. Nor need there necessarily be statable rules. There may be something like phronesis, a skill of knowing how, rather than knowing that. Certainly we can give justifications, but when they reach an end, we can only say “this is what we do” (Wittgenstein 2003, Remarks 151, 179). Normative foundations can be thought of as analogous to dictionaries which have normative force for correct word usage but need not play this authoritative role. When dictionaries are used to settle disagreements, this is not because their authority flows automatically from their correctness, but because we have endorsed their authoritative use for particular situations. The insistence that Foucault articulate normative principles is based on a picture of principles as being necessarily authoritative.

Habermas urges us to endorse the authority we reasonably grant to science, in the form of his foundational criteria which harness developmental sciences. These foundations - the concepts of communicative reason, and the distinctions that support them – are, though fallible, based on the best
knowledge we have. By their universality, they are something we should accept for constructing social theories. If they turn out to be wrong, we will change to what future science shows us is the case. Habermas’ strategy of discovery employs reconstructive sciences to show what ideals are necessarily presupposed for us to be the sort of reasoning and communicative creatures that we are. Foucault’s strategy of self-transformation reminds us, that such presuppositions only play the authoritative role that Habermas wants them to play, if we endorse that role. Foucault wants us to accept responsibility for this endorsement, because he thinks that “the role of the intellectual is not to tell others what to do”. He deplores “all the prophecies, promises, injunctions, and programs that intellectuals have managed to formulate over the last two centuries” (PPC, 265). His genealogies focus our attention on how such universalist histories, anthropological foundations and traditional emancipatory theories have been blind to their own dominating tendencies. Habermas sees Foucault caught in a performative contradiction here. I argue that the performative contradiction lies elsewhere and Foucault carefully steps around it.

Whether or not we share Foucault’s concerns about universalist foundations and theories, we can grant that his commitment to undermining foundations is itself a foundation. We cannot reject normativity in toto. To attempt this would be to adopt a normative stance against norms. Foucault’s lack of commitment to foundations is a commitment to this lack, in fact a foundation. This is not a matter of indifference to Foucault, but something he argues for doggedly, yet obliquely. If Foucault were to simply announce this foundation, he would immediately be embroiled in performative contradictions, much like commanding someone to be free. It is not surprising that we see a certain circumspection and tentative quality in Foucault’s accounts of his project. Foucault is not rejecting, but distancing himself from regions of normativity by suspending normative force. What he is rejecting is the reliance on norms as prescriptions, determined in advance, to orient future judgements and actions. Such prescription would stifle “more secret, more possible freedoms, and more inventions in our future than we can imagine in humanism” (Foucault 1988, 15).

Foucault remains in conformity with his foundation, whilst not relying on it by invoking it as either a test in a decision procedure, or a justification of the form A, therefore B. He requires no normative foundations for these purposes, because his judgements are responses to concrete experiences of oppression, either his own, or those of others with whom we can identify, independently of philosophical justification. The difference between conformity with, and reliance upon, a foundation is a question of where we place ultimate authority. We can open the dictionary and declare “this dictionary agrees with us” or “this is what it really means”. But Foucault’s normative foundation is not of a type that would satisfy Habermas, who wants an explicit normative foundation to play an authoritative role. Rather than leading to any determinate prescription, Foucault’s normativity
emphasises a process, a way of being which embodies an ethic not guided by norms, yet binding him to constantly question norms.

Foucault incites us to think differently whilst not wanting to tell us what or how to think. His inflammatory rhetoric should be understood not as a total condemnation of Western modernity, but a response to its complacent self-assurance. “I always analyse quite precise and localised phenomena ... I don’t say that Western civilisation is a disciplinary civilisation in all its aspects” (Foucault cited in Kelly 1992, 371). Certainly DP contains the rhetorically charged description of the “carceral archipelago”. Most frequently however, Foucault adopts a distanced “agnostic” stance that steps back from judgement. \[137\] If reason is always imbricated with power, convincing people with reasons is not enough. Hadn’t the repressive hypothesis been a matter of people convinced of the need for liberation ending up enslaved to new forms of power/knowledge? Foucault’s principal aim, that he became increasingly clear about, was not to convince, endorse or reject, but in the first instance, to make something into a problematic object, something that we must come to recognise as, in a sense, our own creation and responsibility. Like the gadfly in the agora, Foucault’s project is not to give answers but to prompt us to question and think, to accept responsibility, “to care for ourselves”.

Foucault urges “radical transformation” (EW3 456-7). He seeks “to construct” himself, and to “invite others to share an experience of what we are ... an experience of our modernity in such a way that we might come out transformed” (EW3 242). This experience challenges categorical frameworks, in a reformation of subjectivity analogous to a religious or psychoanalytic experience. Foucault doesn’t explain this. Like the Cynics, his philosophy is a way of life by which the ideal of freedom is shown, not said. \[138\] This non-discursive critical practice takes the form of a “true” life shown by one’s bearing and resolution, in Foucault’s case the embodiment of principled resistance as parrhesia. Saying inevitably projects universal categories. But to show offers an experience. Transformation does not come about through the force of the better argument, reasons justified by foundations, but experiences which change one. \[139\] Foucault emphasises that his histories are “experience-books”, not “truth-books”. It is not truth that is important, but the “experience that the book makes possible” (EW3, 243).

O’Leary discusses two notions of “experience” in Foucault’s work (O’Leary in O’Leary, T. Falzon, C. (eds) 2010). Everyday experiences arise from the frameworks that emerge from large-scale historical changes. Transformative experiences are not determined by such frameworks but transform

\[137\] We see this clearly in his books BC, AK, OT, HS1, HS2, HS3  
\[138\] I follow Wittgenstein’s thought that ethics belongs to that realm where things cannot be said but only shown (Tractatus, 4.1212, 6.421).  
\[139\] In OT, Foucault refers to his analysis as an “experience” of Order. (OT p. xxiii)
us such that we are no longer as before. These notions can’t be separated. It is precisely the investigation of the historicity of experience in the first sense that enables its history in the second sense. The aim of Foucault’s genealogy is to encourage, or even provoke, an ongoing problematisation of the present (EW1, 223-252). It seems that transformative experience occurs on two levels. Firstly we may be oriented to specific political actions by realising that concepts such as “delinquent”, “pervert” and “madman” are invented. This is not to say that such inventions actually limit freedom, but only that we must be alert to this danger. Secondly, a broader awareness opens on realising that all our concepts, social concepts particularly, could be invented.

In his late work, Foucault calls the process of self-transformation an “aesthetics of the self”. Like aesthetics, it is not rule-bound, prescribing how we must respond, but rather opens up further possible responses. It calls for our assent not by legislating, but by claiming, assent. This aesthetic self is not solipsistic, but social and political. Self-transformation entails struggle with the pastoral function of state power that “categorises the individual, marks him by his own individuality, attaches him to his own identity, imposes a law of truth on him which he must recognise and that makes individual subjects.” Political struggles “revolve around the question of who are we? They are a refusal of a scientific or administrative inquisition which determines who one is” (Foucault in Rabinow & Dreyfus 1983, 212). The politics of who we are remains vulnerable to resistance.

For Habermas, a foundation is necessary for committed judgement and action. So what sort of foundations would Habermas require from Foucault? King has noted that in his discourse ethics, Habermas formulates two principles, a universalising principle (U) and a discourse principle (D), that operate as unassailable foundations for moral discourse (King 2009; MCCA, 65-6). I refer to these principles only to illustrate the sort of foundations that Habermas presumably thinks Foucault needs. Habermas claims these principles as foundations because one can only continue questioning beyond them on pain of not making sense. They are necessary, since in denying them one contradicts oneself. Habermas needs such foundations because his primary aim is to safeguard the normative structure of the threatened lifeworld. He assumes that this is what Foucault must have, because he sees Foucault having a “serious intent of getting a science underway” that aspires to true objectivity of knowledge (PDM 279).

To see what Habermas is getting at, we should ask not what science is, but what it does. Within its institutional setting, science enforces boundaries between itself and non-science by recognised conventions, such as positions, organisations, qualifications, journals which ensure the open-ended fallibilistic orientation to truth constitutive of scientific activity. Within the scientific community,
certain fundamental claims are considered hypothetical and open to refutation.\textsuperscript{140} Yet within society more broadly, science takes on an authoritative role, as a set of axioms at which questions should stop. As a practice, science may be an open-ended inquiry, but its functional role as it enters the modern lifeworld is to provide a point on which we can all agree, since its claims have nothing to do with our predispositions as individuals or members of particular communities, but only with the way the world is for all human beings. Certainly what science tells us will change, but whatever science at any particular time tells us, we should accept as a guide to action.

Habermas is suggesting that Foucault is engaging in a discourse which, like science as it enters the lifeworld, draws on foundations which we should accept. By so doing, science offers practical confidence to the lifeworld, which we can act upon, at least provisionally. Without such foundations, we are compelled to keep arguing rather than solving problems by achieving the sort of pragmatic agreements and outcomes that science enables. Habermas’ discourse ethics provides an example of what he might accept as a foundation, what could be called “science” in the broad sense of something authoritative, where questioning must stop, in this case, on pain of contradiction.

Habermas thinks Foucault is “getting science underway” because he assumes that he is what Koopman calls “normatively ambitious”, attempting to discover determinate facts to definitively settle questions (Koopman, 88-98). Habermas presumes that the only way to legitimately orient thinking is to legislate it in terms of striving to reconcile the real and ideal through the rule-bound use of reason. But Foucault’s genealogy is normatively modest, not proving or disproving, supporting or subverting, but alerting us to dangers and encouraging us to be sceptical. Taken as a whole, his genealogy aims not to legislate but to clarify and intensify awareness of present dangers. Modern power isn’t necessarily bad but is problematic and demands serious attention. “This is a domain of very complex relations, which demand infinite reflection” (Foucault 1988a). To make a determinate judgement forecloses reflection. Even Foucault’s sympathetic interpreters, Rabinow and Dreyfus, have difficulty here, arguing that Foucault “owes us a criterion of what makes one kind of danger more dangerous than another” (Rabinow and Dreyfus 1983, 264). The problem is that, since the demise of ancient philosophy’s \textit{care of the self}, the answer to such questions is always conceived in terms of a science of \textit{knowing the self}, humanity or society. Whilst “knowing” suggests a programmatic response to a fixed object, “care” suggests a more hermeneutical attitude of mutual adjustment and responsiveness to the subject’s own purposes, potentials and meanings. Foucault wanted to separate our ethics and lives from knowledge and science.

\textsuperscript{140} This is not to deny that on a synchronic view, “normal science” is constrained by “paradigms”, or broad sets of assumptions which remain beyond question. (Kuhn 1996)
Genealogy’s radical perspectivalism is clearly at odds with the conventional truth claims of science buttressed by reasons and foundations. Foucault wants to show that science’s “value freedom” is itself a value that excludes other values. Genealogy takes its self-referentiality seriously. There is no perspective outside the power-laden discourses from which to speak truth. This doesn’t mean genealogy isn’t true, but rather that truth is not what we thought. The truth of genealogy is not anything like the truth that science pursues. For Foucault this truth is not enough. We need to ask about regimes of power, not to reveal claims as simply untrue, but to uncover how truths have emerged. Whilst Habermas builds his discourse ethics on unassailable, though fallible, foundations – points where questioning must necessarily end - these are precisely the necessities that Foucault would question. It may seem preposterous to challenge a performative contradiction – surely a point beyond which we can make no sense – but Foucault has argued that what can count as making sense is never stable. His histories chart stranger things that our provincial imaginations can barely grasp – the Renaissance world where everything is connected by resemblances, the coming into being and the passing away of Man, the archaic truth regime of the ordeal. These are contingent universals, or as Foucault would say, historical a priori. Truth is not a context-transcendent realm but a thing of this world only grasped in relation to specific constellations of power-knowledge-subjectivity. He is acutely aware of how power creeps in under the cover of hypostatised humanistic universals to bend them to its own ends. Science is complicit because it is never pure, but always embroiled in constellations of power/knowledge.

Foucault is not, as Habermas claims, aspiring to science, whereby a fixed subject seeks to justify determinate conclusions. His critique is not “refusal and denial, but rather an investigative work that consists in suspending as far as possible the normative system which one refers to in order to test and evaluate it” (Foucault, cited in Lemke 2012, 61). By both “suspending” and “referring to” the normative system, Foucault’s critique reflects self-referential modernity’s “rational criticism of rationality” and the imbrications of its moments. It suspends normative frameworks to avoid the danger of assimilation into prevailing constellations of power. It also refers to the normative framework, not by directly evaluating social practices, but bringing this evaluative framework itself into view. The genealogy of forms of rationality follows “the course of a precarious and fragile history” (PPC, 35). This history isn’t merely descriptive. “The history of various forms of rationality is sometimes more effective in unsettling our certitudes and dogmatism than is abstract criticism” (EW3, 323). By “unsettling our certitudes and dogmatism” critique is linked to the space of concrete freedom by the possibility of transformation. Yet genealogical critique doesn’t legislate, but invites and prompts transformation and change (Foucault, 1981, 69-72).

Foucault’s critique questions the relations between knowledge and power that issue in epistemological certainties that buttress ways of organising social relations that foreclose alternative ways. It implies
that these certainties are constructed by forms of power/knowledge and limit the possibility of thinking otherwise. One might reasonably ask what good is “thinking otherwise” if we aren’t guided by principles to gauge success or failure. As Butler points out, there are no reassuring answers to give to such a line of criticism, since reassurance is not the aim of critique (Butler 2001). Foucault’s critique emerges from an existing crisis within the epistemological field in which we already can’t see the way forward. His archaeological analyses emphasised that the categories which order social life produce certain realms of unspeakability. It is from the awareness that dominant discourses have produced these impasses that critique emerges as an exploratory experimental practice.

6 POWER AND REASON

The two moments, respectively thematised by Foucault’s and Habermas’ strategies, can illuminate the substantive tension that “lies at the very heart of social and political theorising … between rationality and power” (Allen, 3,4). Both Habermas and Foucault recognise what McCarthy calls the “impurity of reason … its embeddedness in culture and society, its entanglement with power and interest, the historical variability of its categories and criteria, the embodied sensuous and practically engaged character of its bearers” (McCarthy 1991, 43). However, their different stances have led to suspicions that Foucault collapses truth into power and Habermas insulates truth from power. Rather than consider these arguable suspicions directly, I want to clarify the contrast between Foucault’s and Habermas’ understanding of power and knowledge.

Foucault insists on the inseparability of power and knowledge. In contrast, Habermas draws a strong distinction between power and validity. Conscious of a lifeworld threatened by scientism, technocratic consciousness and systemic imperatives, it is not surprising that he reacts strongly against Foucault’s notion of power/knowledge which concatenates an ensemble of bodies, concepts, discourses, practices, objects and subjects. Rather than the progress of scientific knowledge, Foucault sees constant change brought about by realignments of heterogeneous forces within omnipresent power/knowledge. Habermas objects that “from [Foucault’s] perspective, socialised individuals can only be perceived as exemplars, as standardised products of some discourse formation – as individual copies that are mechanically punched out” (PDM, 293). He thinks that Foucault’s notion of power/knowledge leaves no room for freedom. He sees Foucault collapsing knowledge into all-pervasive power. Habermas would argue that fundamental distinctions such as between validity and power, along with the distinctions we saw in Chapter 7, have been endowed to us by social
evolution. Given their pragmatic role in safeguarding the lifeworld’s communicative resources, it is not clear that we should, or even could, dispense with them.

But Foucault doesn’t reduce knowledge to power or identify power with knowledge (PPC, 43). Whilst knowledge and power are not the same, they are so entwined that we can only analyse them in relationship to each other. Foucault’s “power” is not in opposition to knowledge or freedom but is a condition of possibility which not only constrains, but enables, knowledge and freedom. Whilst Habermas objects to the notion that power is constitutive of knowledge, Foucault suggests that we “abandon the whole tradition that allows us to imagine that knowledge can exist only where the power relations are suspended and that knowledge can develop only outside its injunctions, its demands and interests” (DP, 27). By conceiving power in terms of relations, Foucault considers subjects are free when they can modify, negotiate or reverse these relations in ways that increase capacities and possible modes of thought and action (Foucault in Rabinow & Dreyfus 1983, 221). Freedom for Foucault is not a state one achieves, once and for all, but an “undefined work” of thought, action, self-invention. By framing power in terms of reciprocal relations of “incitation and struggle” between free subjects, Foucault’s power relations can, as Falzon suggests, be seen as dialogical (Falzon, 2010). Within this “dialogue”, individuals are active. But forms of social organisation emerge out of the anonymous interaction of countless struggles, much as we saw in PP. This dialogue gives rise to forms that are not rationally necessary, but contingent, unjustified, surpassing any individual intention. The self is both historically formed by interactions between situated perspectives and can actively resist what is historically given.

Whilst positing a strong categorical distinction between power and validity in terms of the real and the ideal, Habermas recognises that reason is never insulated from power. “The critique of validity claims cannot ultimately be separated from a genetic consideration that issues in an ideology critique – carried out from a third-person perspective – of the mixing of power claims and validity claims” (PDM, 323-4). Whilst Foucault’s genealogy and Habermas’ critique depend on the specific conditions of modernity, Habermas objects that Foucault identifies this context-dependency with context-boundness of validity. Habermas in contrast, deploys criteria of rationality which transcend any context, thus making critical evaluations dependent on an ideal perspective that all would agree to. Foucault would see this regulative ideal as not something waiting to be discovered, underlying communication, but an invention. We cannot “dissolve” power relations “in a utopia of perfectly transparent communication” (FL, 447). Since there is no “outside” of power, Habermas’ regulative ideal denies the possibility of the only kind of freedom we have - concrete freedom within games of power/knowledge. Rather than making judgements in terms of an ideal absence of power, Foucault

---

141 For example, distinctions between doing and occurring, or made and grown, (see Chapter 7)
thinks we should endeavour to free social practices from “domination”, in which asymmetrical power relations remain frozen. By endorsing its authoritative role, Habermas exempts his “utopia” from “the rational critique of rationality”. Foucault’s *ethos* doesn’t seek our endorsement for an authoritative role, but presents itself as a particular finite standpoint. Open to the rational critique of rationality, it can self-reflexively consider the costs and possibilities it might blind us to. Rather than eliminating power, by invoking an ideal outside power, we should, by the practices of the self, allow “games of power to be played with a minimum of domination.” (FL, 447)

We see here the play of imbrications of participant and observer perspectives. As *participants* in contexts of action, we orient ourselves to unproblematic truth that we regard as unconditioned by, and beyond, any actual justification. Habermas, the participant in the game of discovery, seeks something that whilst fallible, is sufficiently secure to commit to. Yet by suspending commitment, this game may be made the object of third-person accounts of its contexts and conditions, its regime of truth. Foucault, adopting this observer’s perspective, is of course imbricated as a participant in his own game of observer. In generating beliefs about beliefs, he also commits, as a believing participant, to this observer’s perspective, *directed towards* the object beliefs, in which he suspends commitment. There is no point at which the movement of stepping back to assume the other, already imbricated, perspective must end. Habermas insists on a threshold, beyond which one must overcome the paralysis of deliberation in order to act or commit to a determinate judgement. Foucault maintains this tension which cannot be resolved in favour of either perspective.

If we view Habermas’ quasi-transcendentalism in terms of an ongoing process, not simply providing a regulative idea to legislate the lawful use of reason, but itself open to the “rational criticism of rationality”, the apparent tension between power and validity appears more productive than oppositional.\(^{142}\) The actual meaning and force of communicative norms is always to some degree contextualised within actual processes of communication. We can thus respond to critics who see his idealised presuppositions of communication, as authoritarian rules that demand adherence from all. Habermas’ idealisations are better seen as ongoing appeals to future communities which they may endorse as they interpret them within new contexts. There is no need for a universalist assertion of validity, already warranted in advance of its interpretation into those contexts (see Hoy and McCarthy 1994, 72). Without this *proviso*, Habermas seems caught in his own performative contradiction, legislating how future communities are to be free and reasonable. To be applied, the bare formality of these idealisations must be fleshed out within actual contexts that require interpretation and additional assumptions to provide content. Given its formal and procedural character, Habermas’ idea of society based on uncoerced agreements, arrived at in free and equal exchanges, need not be the depiction of a

\(^{142}\) This is an approach adopted by various commentators including McCarthy (1991), Koopman (2013).
concrete utopia as Foucault assumes (Foucault 1987, 18). The pragmatic force of Habermas’ context-
transcendence is that its regulative dimension opens reason to self-correction, by enabling a foothold
for criticism.

Habermas’ categorical distinction between power and validity is however, more problematic in the
form it takes as the system/lifeworld distinction in his two-level concept of society. The distinction is
suited to a mode of analysis directed at addressing the problems of modernity – rampant scientism,
bureaucracy, technological consciousness and capitalism - in the spirit of Weber, Horkheimer and
Adorno, updated to account for the phenomena of late capitalism. Habermas also had to accommodate
concerns he had raised earlier in his criticisms of systems theory as a negative utopia of technological
control over history which totally objectified society (TRS, 104). However, he remained drawn to
systems theory as a means to capture aspects of power that bypass everyday consciousness. TCA’s
solution is to conceptually partition the lifeworld as the home of communicative reason where the full
scope of validity spheres is operative from systems, which harbour power that infiltrates and distorts
the lifeworld’s communicative resources. McCarthy thinks that this account “cedes too much territory
to systems theory” (McCarthy 1991, 153). As we saw in Chapter 6, the partitioning comes under
pressure at certain points, for example, in Habermas’ account of bureaucratic organisations.

Following the critical reception of TCA, Habermas considerably softened the lifeworld/system
distinction, downgrading the colonisation thesis to an empirical question of the direction of
colonisation (Honneth & Joas, 254). However, whilst not denying the operation of power within the
lifeworld, his primary focus has remained power as a threat to the modern rationalised lifeworld from
autonomous systems which have broken free. Within the modern lifeworld, systems bypass
communication oriented to uncoerced mutual understanding by employing steering media such as
money or power to coordinate outcomes of strategic action. Habermas basic argument is that these
two components – the lifeworld’s communicative resources and systems steering mechanisms - are
out of balance. The problem with this account is that it doesn’t adequately explain how power,
originating in and endemic to the modern lifeworld bypasses rational scrutiny. Habermas doesn’t
adequately consider subjection as a mechanism for reproduction of forms of social subordination, for
example, racist or gender stereotypes perpetuated through lifeworld socialisation, clearly forms of
domination not adequately captured by the account of colonisation by systemic imperatives. Such
difficulties don’t arise only from the direction of influence between systems and lifeworlds, but call
into question the theoretical efficacy of the distinction itself. The distinction clearly enables an
analysis of the problems of modern society. Habermas’ fallible social theory fruitfully bears on the
dysfunctional state of modern democracies, giving us concepts such as colonisation of the lifeworld,
fragmentation of consciousness, systematically distorted communication and the means to analyse,
diagnose, model and interrogate society. However, the cost of the lifeworld/system distinction is seen in what is lost by the imposition of categories that are actually inextricably compounded.

A further difficulty is the depoliticisation of scientific technology by Habermas “rebalancing” picture which implies that the logic of “labour”, natural science and technology, whilst requiring curtailment to its proper domain, appears neutral “in itself”. I want to complicate this picture by returning to Feenberg’s argument that technology is a media like power and money, which steers actions strategically (Feenberg 1999, 166-72). Whilst technology doesn’t communicative burdens in the manner of power or money, it unavoidably addresses all dimensions of validity. Technological objects are not purely functions but also signs, not only signifying their functions, but a vast range of social values, for example, status or power. And whilst both functioning and signifying, they also physically induce certain social arrangements non-consensually. Not only the panopticon, but production lines, work spaces, electronic social media all prescribe certain social arrangements and preclude others. But whilst technology in this way co-ordinates action as a systemic medium, it equally belongs to the lifeworld. As an amalgam of function, signification and social construction, technology is as much open to the lifeworld’s contestation, as it is to systemic employment that bypasses contestation. One can for example, either be shaped by the impoverished communication that the pre-digested format of social media forces upon users, or argue against that impoverishment. Given this ambivalence of technological artefacts, the lifeworld/system partition may conceal as much as it reveals.

This loss is also seen when we consider juridification, the use of law as a medium that non-consensually moulds behaviour. I have discussed the important role that scientific discourses play in this process by providing authoritative justifications which secure consensus and avoid conflict. This enables a media-like steering of society, whereby citizens passively accept science, which remains inaccessible and “split-off”. But Habermas also tells us that expert discourses can be “ultimately” redeemable. Habermas’ prevarication on the question of expert scientific discourses as being, on one hand, subject to “condensation” whereby we put reasonable trust in a professional person, and on the other hand, being “split-off”, whereby that trust comes under suspicion, suggests that this picture of system versus lifeworld, with a movement of power or influence one way or the other, prevents us from seeing the complexity of interactions falling outside Habermas’ binary schema.

Whilst not suggesting that some form of functional perspective is valuable in getting at the counter-intuitive aspects of society not visible from the lifeworld, we can reasonably ask why Habermas incorporates systems theory to generate the theoretical structures he does. It appears that he is drawn to systems theory to enable a comprehensive and robust theory of society. Habermas wants to overcome the pessimism of his Frankfurt predecessors by establishing a firmer normative basis. He enlists the sciences to underpin a progressive developmental account yielding the presuppositions
necessary for our communicative practices as criteria for this development. The distinctions and idealised presuppositions on which this account rests, provide a fallible foundation that gives as much certainty as possible to orient action. Habermas is able to account for the dominance of instrumental reason in more concrete terms than his predecessors, specifically in the increasing influence of the economic and administrative systems. He is acutely aware of the irony of modernisation in which “the communicative potential of reason has been simultaneously developed and distorted” (PDM 315) He highlights the ambivalence of modern rationalisation that initially released communicative potentials, only to undermine those potentials. Although perceived by post-modernist critics as authoritarian, Habermas’ idealisations and use of reconstructive sciences are much more a standard by which to recognise and criticise dogmatic authoritarianism.\(^{143}\) Firstly, his bases are fallible, subject to empirical corrections and secondly, every application will require interpretation, and here philosophy is drawn in as interpreter between the lifeworld and the specialised sciences. Habermas’ strategy of discovery, harnessing science from a vantage point outside the everyday lifeworld, offers determinate tools to not only diagnose, but go some way in prescribing remedies for the ills of modernity.

Foucault’s dispersed power/knowledge provides none of these tools, and nothing like a prescription. In contrast to Habermas, his broad categories of analysis lead to a more abstract account of the truncation of reason by Enlightenment humanism, the normalising effects of disciplines and biopower. This alerts us to the dangers of frozen thought, captured by constellations of power/knowledge and drawn into contexts where it limits freedom. Foucault’s strategy of self-transformation provides a vantage not only outside the everyday lifeworld, but outside science which he problematises, and truth which he relativises. This is not to say that he doesn’t believe or value science or truth. His project is to historicize the transcendental, the necessary, the universal wherever he meets them, to show how much of our world is our “invention”, how the most secure and obvious categories of scientific thought depend on us, and would be different if we were different. Foucault would want to show that what Habermas has discovered, has also been “invented”. He would emphasise the porosity of Habermas’ categories and the negotiability of the moments of freedom and resistance. In nominalist vein, his radical historicism would blur Habermas’ categories, to draw out what fails to fit them as complex and hybrid.

Foucault would criticise Habermas’s theoretical framework not only for its abstractions, but the universalism of the “universal intellectual” tradition that derives from the jurist and juridical tradition of the West (FR, 70). The problem is that “certain great themes such as humanism can be used for any purpose whatsoever” (FR, 374). Whilst Foucault himself heuristically employed general frameworks, he explicitly avoids the authoritative claims of theory and provides a genealogical account of science

\(^{143}\) Lyotard (1988) and Cooke (2001) criticise Habermas in terms of authoritarianism.
to highlight its constructed contingent nature. Rather than seeking standards to be applied universally, he dramatises excesses of power in particular situations. He acknowledges the cost of this commitment to local partial inquiry is to “give up ever acceding to a point of view that could give us access to any complete and definitive knowledge of what may constitute our historical limits” (PT 115). Foucault’s genealogies raise the possibility that claims of context-transcending validity, by obscuring their entanglement with power, originating from particular conceptual orders, collude in repression. For Foucault, the risk is that Habermas’ structures, however apt they seem to the current situation, can be drawn into constellations of power and, granted the authoritative backing of science, become a universal force projecting across all contexts.

It would be difficult to reconcile Foucault’s notion of power/knowledge with Habermas social-theoretical framework of lifeworld/system, buttressed as it is by a series of distinctions between formal and empirical levels - logic vs dynamics, universal moral norms vs concrete ethical life – which serve to maintain Habermas’ robust conception of autonomy. But this doesn’t mean they are opposed. My focus has been to highlight the motives behind the two strategies to reveal them as moments within the broader frame of modern self-reflective thought. The two strategies are reflected not only in Foucault’s and Habermas’ employment of science, but their respective accounts of science and its development. On one hand there is Foucault’s contingent, anarchic, impersonal power/knowledge, on the other, Habermas’ progressive science reflecting the developmental logic of communicative action.

7 SCIENCE, RELIGION AND PHILOSOPHY

Foucault’s genealogies chart the course of concepts and discourses on their way to becoming sciences, as they interact with each other and the non-conceptual and non-discursive under different truth regimes. His histories abstract from the subject, to reveal the play of power/knowledge, and from truth, to deal only with the contingencies of what is taken-as-true, so revealing the fine-grain of events beyond consciousness. In chapter 4, we saw Foucault account for the emergence and dominance of a particular form of truth, the demonstrative scientific truth of the “Cartesian moment” (HM, 550-74). In chapter 3 we saw Foucault sketch the history of a particular form of truth in his genealogical account of the empirical sciences emerging from judicial techniques of inquiry (EW3, 4; PP, 246; DP, 225). The human sciences grow out of a completely different technology of power, which becomes pervasive in modernity - the examination, linked to observation and normalising judgement. They form regions, modelled on the empirical sciences and, whilst initially connected to practices of power in asylums, schools, prisons and hospitals, are generalised throughout society. Medicine establishes the normative basis for the entire social sciences. Human and natural sciences are linked to other
scientific and non-scientific discourses and practices such as law and morality, to constitute circuits of power.

In his 1979 lectures Foucault again relativises scientific truth in relation to the broader frame provided by the concept of “regimes to truth”. Rather than accept “truth” as an obvious good to be sought at all costs, Foucault opens up the question of the limits to truth and its technologies of power. His genealogical accounts, charting the emergence of areas of knowledge and their inextricable links to power, are not intended to be merely descriptive, but transformative and critical. Past *epistemes* throw our current situation into sharp relief. Unitary knowledge is broken apart and what is taken as natural, universal and necessary reveals itself to be contingently constructed. Foucault doesn’t intend to convey a story of either progress or regress. Nor does he find anything to point the way to progress. But we may be transformed by an experience that opens up possibilities for changing social practices and relations, although not in any way that Foucault can specify in advance. He wants us to feel the contingency of our present to be as strange as earlier *epistemes* from which it accidently developed. His densely factual histories of discrete areas of the human sciences - sex, mental illness, criminality - refuse systematic theories.

Habermas provides a progressive developmental account of science. Consistent with the Frankfurt School approach, he harnesses science in his philosophising. Empirically plausible universalistic theories are employed to bear on contemporary crises, bringing into view the social evolutionary phenomena of a progression of developmental milestones that provide a normative standpoint towards modernity. Whilst the sciences have their risks and setbacks, there is a logic to their developing the way they do. This “logic” is firstly (in KHI) a logic of interests. Here the empirical sciences are guided by a technical interest in dealing with an objective world, whilst the human sciences are guided by a practical interest in the social world. In his mature theory, this logic is vested in the development of communicative reason which makes increasingly finer distinctions between objective, inter-subjective and subjective worlds, separating out what is in the world from what is in us and between us. Communicative action propels us forward. The main danger, whether in Habermas’ earlier Kantian theory of KHI, or the communicative reason of TCA, is one of balance, whereby this development is skewed by the contingencies of power which engender a limited form of reason, taken as reason *per se*. This view of science’s developmental logic and the threats of truncated reason forged by its relationship to power is consistent throughout Habermas work, from his early critiques of positivism to his colonisation thesis and his recent defence of the lifeworld structures against reductive naturalism.

In Nietzschean spirit, Foucault would insist that whilst social evolution has occurred, this need not be viewed as an improvement. Certainly we have become more able to distinguish validity claims, but is
this better? Like Habermas’ use of science, genealogy also enables one to question what was
previously unquestionable. Yet Habermas wants conclusions, and so draws a line at what must be
necessarily presupposed. Foucault questions what seems to be necessarily presupposed, as being
presupposed only by creatures such as Man. Habermas’ worry is that Foucault’s critique stops
nowhere, is totalising and loops around to pull the rug from under its own feet. Foucault’s worry is
that Habermas’ theory nails the rug to the floor so as to exempt itself from modernity’s “rational
critique of rationality”.

We must also consider non-science, that which is contrasted to science or, in positivistic terms, that
which science overcomes - religion. Neither Foucault nor Habermas are dismissive of religion like the
positivists, assuming that religion will fade away under the clear light of an expanding scientific
worldview. Certainly Habermas affirms the priority of science and secular discourse. Secular
discourse informed by science not only is, but should be, the lingua franca of modernity. Science
represents a rational advance by its exclusive focus on the differentiated validity sphere of truth to
grapple with the external objective world. Scientific discourses are unproblematic if they don’t
overstep this domain to be drawn into constellations of power where their authority is co-opted to
legislate within other validity spheres. Secular language informed by science, aspiring to transcend
particular perspectives by granting communicative action its fullest scope must have prime place in
politics. However, Habermas is acutely aware that “religious language is the bearer of semantic
content that is inspiring and even indispensible” and that its absence leaves a loss which it seems we
can only lament. He refers to the funeral ceremony for a friend, “an agnostic who rejected any
profession of faith” yet who, by his choice of a church as the place of ceremony, had “publicly
declared that the enlightened modern age has failed to find a suitable replacement for a religious way
of coping with the final rite de passage which brings life to a close” (Habermas et al 2010, 15).
Habermas thinks we should attempt to translate the residue of meaning locked in religious language
into secular language. He acknowledges that at times this eliminates the substance intended and leaves
irritations (FHN, 110). Habermas wants non-believers to shoulder equal responsibility for undertaking
“cooperative” translation of religious to secular language (BNR, 131). Whilst not a straight-forward
task, translation enables divergent traditions to find a common language other than coercive forms of
economic, political or social power.

If Habermas’ view is that of the anthropologist from post-metaphysical modernity, committed to
norms by which science is granted an authoritative position in understanding the world and the respect
of religion’s semantic potential, Foucault is the anthropologist from another planet, looking at the
species from such a bemused distance that both science and religion appear as different regimes of
power/knowledge. In PP he accounts for how the Christian confession migrated into the
psychoanalytic confession, and the fear of monsters was transmuted into the search for instincts.
Foucault’s interest in religion is not as a system of metaphysics or a form of knowledge but as a way of life, a spiritual activity by which one cares for oneself—“spiritual” in a specific sense not involving commitments to religious concepts but rather “the search, practice and experience through which the subject carries out the necessary transformations on himself in order to have access to the truth” (HS, 15). Unlike science where the truth is open to all, the spiritual subject’s access to truth is grounded on ascesis, on transforming her mode of being a subject, of seeing and inhabiting the world, by certain exercises amounting to a sort of conversion (HS, 15). This notion of a critical-experimental practice of self-transformation is essentially political. “There is no first or final point of resistance to political power other than in the relationship one has to oneself” (HS, 252).

In these two responses to religion, we again see the two biographies. Foucault wants to alert us to the dangers of authoritarianism he saw within science. We see him particularly concerned, for example, about the psychiatric power that tells us what we are and must be. To that end, his more radical observer’s view seeks self-transformation in a series of overcomings of whatever appears necessary. Habermas wants a degree of certainty to identify and call out the deceptions and power imbalances on which the authoritarianism of his teenage years fed. He believes that science, by its “external” stance gives us at least some purchase on our situation. Its counter-intuitive perspective infuses the traditional lifeworld with fresh perspectives that aspire to universality, a role that modernity cannot leave to religion. Yet he is all too aware of the imperialistic tendencies of science that, whilst compelling to the modern mind, occlude other dimensions of validity and meaning, some of which reside uniquely in religion. It is philosophy that must traverse these disparate domains to make science at home in the modern lifeworld whilst acknowledging and drawing on religion’s semantic potentials. Philosophy can’t view the human condition from an Archimedean perspective. Whilst drawing on science as a resource providing an “external” view, fundamental questions require philosophy’s capacity to step back from the empirical data to draw it into a broader framework. What Foucault and Habermas both value about philosophy is its anarchic nature, unconstrained by the scientific method. Like Habermas, Foucault thinks that philosophy must find its place within the particular social and historical formation of modernity by attempting to understand, diagnose and change the present by a critique of modernity’s paradigmatic form of reason, science.

Habermas draws on the authority of science to buttress a progressive account of reason, yielding an explanatorily powerful diagnosis of modernity as well as conceptual tools of analysis and prescription. His account of the development of science stresses reason’s autonomy. Despite constant entanglement with power, reason tentatively traces a path of progress towards enlightenment. His fallibilism admits that the world can always be some way that our descriptions fail to capture, thus leaving room for his foundational principles to be found to be incorrect by empirical science which makes progress by ongoing corrections.
Foucault suspends science’s authority, adopts agnosticism towards its truth, so bringing it into view as one form of knowledge amongst others. Reason is not autonomous but is tied to objects like the panopticon, power and bodies. He suspends not only truth, but the “truth regime”, or pre-conditions for what can be taken as true or false. Whilst disconcerting, this stance brings to awareness the unseen endorsement granted to science as the most authoritative knowledge, an endorsement that doesn’t automatically flow from its purported correctness. This revelation of our constitutive activity reminds us of our epistemic responsibilities and the dangers of science and reason being drawn into circuits of power. It follows that what we take for truth shifts not only in response to scientific discovery, but in response to shifts in subjectivity. The former corresponds to what Foucault has called “knowledge of the self”, or what Habermas wants to discover, the latter “the care of the self”, or Foucault’s self-transformation, undertaken by *ascesis*.

A common *telos* is seen in the strand of the Enlightenment tradition which culminates in critical theory, aimed at both understanding and changing society, by liberating human beings from entrapment in systems of dependence or domination, both internal and external. For both thinkers this is a political project in which they seek to change consciousness by revealing it to be conditioned by a vast range of historical and social factors. Both recognise the need for a more reflexive perspective and see philosophy as able to articulate social problems not visible from the specialised perspectives of science which must be seen to have a limited scope. We have seen the restless movement of modernity’s self-referential reason between the imbricated moments of participant and observer. Clearly we need to both posit context-transcending ideals and unmask their status as illusions rooted in power-laden contexts. We need a framework to protect freedoms from anonymous systems of surveillance and discipline. Yet freedom is already lost if the meaning of this framework is beyond question. To this end, Habermas attempts to discover some solid ground to build the modern project out of itself and Foucault urges an ongoing process of self-transformation of thought and action. Science, modernity’s most authoritative form of knowledge, sits on the cusp.
ABBREVIATIONS

FOUCAULT

Books
AK The Archaeology of Knowledge
BC The Birth of the Clinic
DP Discipline and Punish
HM The History of Madness
HS1 History of Sexuality vol. 1
HS2 History of Sexuality vol. 2
HS3 History of Sexuality vol. 3
OT The Order of Things

Collections of interviews, essays, talks.
EW1 Essential Works volume 1
EW2 Essential Works volume 2
EW3 Essential Works volume 3
FR The Foucault Reader
FL Foucault Live
PK Power/Knowledge
PT The Politics of Truth
PPC Michel Foucault: Politics, Philosophy Culture see Foucault, M. 1988(b)

Lectures at College de France:
AB Abnormal 1974-5
BB The Birth of Biopolitics 1978-79
CT The Courage of Truth 1983-84
GOL On the Government of the Living 1979-80
GSO The Government of Self and Others 1982-83
HS The Hermeneutics of the Subject 1981-82
PP Psychiatric Power 1973-74
SD Society must be defended 1975-76
STP Security, Territory and Population 1977-78
WK Lectures on the Will to Know 1970-71
ABBREVIATIONS

HABERMAS

Books
BNR Between Naturalism and Religion
CES Communication and the Evolution of Society
FHN The Future of Human Nature
KHI Knowledge and Human Interests
LC Legitimation Crisis
LSS On the Logic of the Social Sciences
MCCA Moral Consciousness and Communicative Action
OPC On the Pragmatics of Communication
OPSI On the Pragmatics of Social Interaction
PD The Positivist Dispute in German Sociology
PDM The Philosophical Discourse of Modernity
PKHI Knowledge and Human Interests Postscript
PM Postmetaphysical Thinking
PMT Post-metaphysical Thinking
RR Religion and Rationality
STPS The Structural Transformation of the Public Sphere
TCA1 The Theory of Communicative Action vol 1
TCA2 The Theory of Communicative Action vol 2
TJ Truth and Justification
TP Theory and Practice
TRS Towards a Rational Society
BIBLIOGRAPHY

Books by Foucault (by date of original French publication)


Foucault, M. 1979 (1976) *History of Sexuality vol. 1* The Will to Know Penguin Middlesex UK


Foucault interviews, lectures, essays


Foucault, M. 1983 *Fearless Speech* 6 lectures at Berkeley online audio resource http://www.lib.berkeley.edu/MRC/foucault/parrhesia.html


Foucault, M. 1988 “Truth Power and Self: an Interview with Michel Foucault” in Luther H. Martin, Huck Gutman and Patrick H. Hutton (eds.) *Technologies of the Self* University of Massachusetts Press Amherst MA


Foucault, M. 1989a *Foucault Live, Collected Interviews, 1961-84* S. Lotringer (ed) Semiotext(e) New York


**Foucault Lectures at College de France**


**Books by Habermas**


Habermas, J. 1972 (1968) *Knowledge and Human Interests* Beacon Paperback USA


Habermas interviews, lectures, essays


Habermas J. & Luhmann N. 1971 Theorie der Gesellschaft oder Socialtechnologie Frankfurt

Habermas, J. 1975 A Postscript to Knowledge and Human Interests Philosophy of the Social Sciences 3 (1975) pp.157-189


Habermas, J. 1983 “The German Idealism of the Jewish Philosophers” in Philosophical-Political Profiles, Cambridge, MA


Habermas, J. 1992 “Taking Aim at the Heart of the Present” in Kelly (ed)1992, 149-154


Habermas, J. 2007 “The Language Game of Responsible Agency and the Problem of Free Will” in Philosophical Explorations, vol. 10 no. 1 March 2007

Habermas, J. 2007a Interview https://www.youtube.com/watch?v=jBl6ALNh18Q
LITERATURE CITED


Agger, B 1992 The Discourse of Domination: from the Frankfurt School to postmodernism Northwestern University Press Illinois

Alcoff, L.. 2013 “Foucault’s Normative epistemology” in A Companion to Foucault Blackwell Publishing Ltd. C. Falzon, T. O’Leary and J. Sawicki (eds.)


Allen, A. 2013 “Power and the Subject” in A Companion to Foucault Blackwell Publishing Ltd. C. Falzon, T. O’Leary and J. Sawicki (eds.)


Bernstein, R. 2010 The Pragmatic Turn Polity Press Cambridge UK


Boothroyd 1996 “Foucault’s Alimentary Philosophy: Care of the Self and Responsibility for the Other” Man and World 29, 361-86

228


Bracken, P; Thomas, P. 2010 “From Szasz to Foucault: On the Role of Critical Psychiatry” Philosophy, Psychiatry, & Psychology, Volume 17, Number 3, September 2010, pp. 219-228 John Hopkins University Press


Canguilhem G. 1991 The Normal and the Pathological Zone Books New York

Canguilhem, G. 1994 A Vital Rationalist: Selected Writing from George Canguilhem Zone Books New York

Chimisso, C 2003 “The tribunal of philosophy and its norms: history and philosophy in Georges Canguilhem’s historical epistemology” in Studies in History and Philosophy of Biological and Biomedical Sciences 34 (2003) 297-327


Cook, D. 2001 “Two Faces of Liberal Democracy” in Philosophy Today Spring 2001


Daston L. 2000 Biographies of Scientific Objects Chicago University Press Chicago


Falzon, C. 2013 “Making History” in *A Companion to Foucault* Blackwell Publishing Ltd. C. Falzon, T. O’Leary and J. Sawicki (eds.)


Feenberg, A. 1999 *Questioning Technology* Routledge New York


Gilligan, C. 1982 *In a Different Voice* Harvard University Press Cambridge


Gutting, G. 1989 *Michel Foucault’s Archaeology of Knowledge* Cambridge University Press Cambridge


Hacking, I. 2002 *Historical Ontology* Harvard University Press Cambridge Mass

Hadot, P. 1995 *Philosophy as a way of Life* Blackwell Publishing Oxford U.K.

Han, B. 2005 “Is Early Foucault A Historian? History, history and the Analytic of Finitude.” *Philosophy and Social Criticism*. Vol. 31, nos.5-6 pp.586-608


Han, B. 2002 *Foucault’s Critical Project: Between the Transcendental and the Historical*. Stanford University Press Stanford 2002

Hanson, N. 1958 *Patterns of Discovery: An Inquiry into the Conceptual Foundations of Science*. Cambridge University Press.


Heidegger, M. 1977 *The Question Concerning Technology and Other Essays* Harper Row


http://discovery.ucl.ac.uk/1381747/

Ingram, D. 2010 Habermas: Introduction and Analysis Cornell University Press


Kant, I. 1987 Critique of Judgement (trans. Pluhar, W.S.) Hackett, Indianapolis


Kelly M.G.E. 2009 The Political Philosophy of Michel Foucault Routledge Oxon UK

Kelly M.G.E. 2013a Foucault’s History of Sexuality Volume 1, The Will to Knowledge Edinburgh Philosophical Guides Edinburgh University Press

Kelly M.G.E. 2013b “Foucault, Subjectivity, and Technolgies of the Sel” in A Companion to Foucault Blackwell Publishing Ltd. C. Falzon, T. O’Leary and J. Sawicki (eds.)

Kellner, D. 1984 Herbert Marcuse and the Crisis of Marxism University of California Press Berkeley

King, M. 2009 “Clarifying the Foucault-Habermas debate: Morality, ethics and ‘normative foundations’” in Philosophy & Social Criticism vol 35 no. 3 pp. 287-314
Kolakowski, L. 1972 *Positivist Philosophy: From Hume to the Vienna* Circle Pelican Books USA

Koopman, C. 2002 “Revising Foucault: The history and critique of modernity” in *Philosophy and Social Criticism* 36(5) 545-565

Koopman, C. 2013 *Genealogy as Critique: Foucault and the Problems of Modernity* Indiana University Press Indiana

Kripke, S. 1980 *Naming and Necessity* Harvard University Press Cambridge Mass


Lemke T. 2012 *Foucault, Governmentality and Critique* Paradigm Publishers Boulder Colorado


McGushin, E. 2007 *Foucault’s Askesis* Northwestern University Press


Maniglier, P. 2013 “The Order of Things” in *A Companion to Foucault* Blackwell Publishing Ltd. C. Falzon, T. O’Leary and J. Sawicki (eds.)

Marcuse, H. 1964 *One Dimensional Man* Sphere Books Great Britain


Oksala, J. 2005 *Foucault on Freedom* Cambridge University Press


Rabinow, P and Dreyfus, H. 1983 *Michel Foucault Beyond Structuralism and Hermeneutics* (2nd ed) University of Chicago Press Chicago


Sellars W. 1963 “Philosophy and the Scientific Image of Man” in *Science, Perception, and Reality*
New York: Humanities Press.

Smart, B. 2002 *Michel Foucault: Revised Edition* Routledge, New York


Veyne P. 2010 *Foucault: His Thought, his Character* Polity Press Cambridge


Webb, D. 2013 Foucault's Archaeology: Science and Transformation, Edinburgh University Press,