

# **Emergent Representations: Dialectical Materialism and the Philosophy of Mind**

**Joe Faith**

## **Abstract**

This thesis is the result of a simple conviction: if minds are the products of brains (and they are) then surely the study of brains will have something to tell us about the philosophy of mind? Moreover, if minds are also the products of natural and social history (and they are) then surely the study of natural and social history will have something to tell us about mind too? Of course I am not the first to suggest these lines of inquiry but they have both traditionally lead to the same uncomfortable conclusion, namely a virulent scepticism about our ability to know the world. We seem to be stuck in a Faustian bargain in which we can only gain scientific knowledge at the expense of philosophical doubt.

This thesis is an attempt to break this bargain, in which I start from the conviction that we *can* know the world, and then ask what kind of science, both natural and social, can make sense of this ability. We do not just need a philosophy of mind that fits our science, we also need a science that fits our philosophy of mind. We must fiddle with both sides of the equation in order to get a fit. In the course of this fiddling I challenge reductionist and empiricist assumptions about science, I question the philosophical tradition that dates back to Frege's 'linguistic turn', and I draw parallels between Marx's theory of history and Darwin's theory of natural selection. The result is a realist philosophy of mind that is built on our ability to interact with and change the world, rather than on our ability to contemplate it passively.

Submitted for the degree of D. Phil.

University of Sussex

March 2000

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
	<b>PART I: MATTER</b>	
<b>2</b>	<b>Anti-Reductionism</b>	<b>6</b>
2.1	Reductionism and Materialism . . . . .	9
2.2	Anti-Reductive Materialism . . . . .	11
2.3	Downwards Causation . . . . .	14
2.4	Conclusion . . . . .	17
<b>3</b>	<b>Naturalisation</b>	<b>18</b>
3.1	Descriptions and Biases . . . . .	19
3.2	Naturalisation . . . . .	21
3.3	Theoretical Terms, Dispositions, and Causal Explanation . . . . .	23
3.4	Laws and Exceptions . . . . .	27
3.5	Prediction and Induction . . . . .	29
3.6	Conclusion . . . . .	31
	<b>PART II: MIND</b>	
<b>4</b>	<b>Brains and Behaviour</b>	<b>33</b>
4.1	Neuropsychology and Neuroethology . . . . .	33
4.2	Representation and Explanation . . . . .	37
4.3	South Coast AI . . . . .	39
4.4	Conclusion . . . . .	45
<b>5</b>	<b>Intentionality: Insides</b>	<b>46</b>
5.1	Opening the Black Box . . . . .	46
5.2	Anti-Turing . . . . .	48
5.3	Externalism . . . . .	52
5.3.1	Epistemological Externalism . . . . .	53
5.3.2	Metaphysical Externalism . . . . .	54
5.3.3	Brains-In-Vats . . . . .	58
5.4	Emergent Representation . . . . .	59
<b>6</b>	<b>Intentionality: Outsides</b>	<b>61</b>
6.1	Sense and Reference . . . . .	61
6.2	Non-Conceptual Content . . . . .	65

6.3	Affordances and Objects . . . . .	68
6.4	Conclusion . . . . .	70

**PART III: NATURAL HISTORY**

<b>7</b>	<b>Functions and Norms</b>	<b>73</b>
7.1	Functional Explanation and Darwinian Norms . . . . .	73
7.2	The Function of ‘Function’ . . . . .	75
7.3	The Function of Behaviour . . . . .	77
7.4	Conclusion . . . . .	78
<b>8</b>	<b>The Role of Genes in Natural Selection</b>	<b>79</b>
8.1	Evolution and Mendelian Inheritance . . . . .	80
8.2	Nature and Nurture . . . . .	82
8.3	Inheritance and Mechanism . . . . .	86
8.4	Evolution and Development . . . . .	88
8.5	Conclusion . . . . .	90
<b>9</b>	<b>The Role of Vehicles in Natural Selection</b>	<b>91</b>
9.1	Burying Vehicles . . . . .	92
9.2	Counting Genes . . . . .	94
9.3	Counting Replicators . . . . .	97
9.4	Fitness . . . . .	101
9.5	Conclusion . . . . .	102

**PART IV: SOCIAL HISTORY**

<b>10</b>	<b>Social Evolution</b>	<b>105</b>
10.1	Natural and Social History . . . . .	105
10.2	Memes and Vehicles . . . . .	107
10.3	Memes and Power . . . . .	110
10.4	Marx’s Theory of History . . . . .	112
10.5	Memes and Symbols . . . . .	116
10.6	Lamarckian Inheritance and Signalling . . . . .	119
10.7	Memes and Adaptation . . . . .	121
10.8	Conclusion . . . . .	122
<b>11</b>	<b>The Good, The True, The Beautiful</b>	<b>124</b>
11.1	Truth and Success . . . . .	125
11.2	Scientific Objectivism . . . . .	128
11.3	Ethical Relativism . . . . .	130
11.4	Conclusion . . . . .	133
	<b>Bibliography</b>	<b>135</b>

# Chapter 1

## Introduction

---

Science explained people, but could not understand them. After long centuries among the bones and muscles it might be advancing to knowledge of the nerves, but this would never give understanding.

— E.M.Forster, *Howards End*

**This thesis is the result of a simple conviction: if minds are the products of brains (and they are) then surely the study of brains will have something to tell us about the philosophy of mind? Moreover, if minds are also the products of natural and social history (and they are) then surely the study of natural and social history will have something to tell us about mind too? Of course I am not the first to suggest these lines of inquiry but they have both traditionally lead to the same uncomfortable conclusion, namely a virulent scepticism about our ability to know the world. We seem to be stuck in a Faustian bargain in which we can only gain scientific knowledge at the expense of philosophical doubt.**

The first way to reach the sceptic's conclusion is *via* neuroscience. If our thoughts are somehow produced by neurological events inside our heads then, as Descartes argued, our knowledge of the contents of our own minds will inevitably be more certain than our knowledge of the outside world distilled from sense-data. The modern inheritors of this tradition are the children of the 'linguistic turn' who hold that thought is primarily a matter of internal processing, though those internal processes are now taken to be linguistic rather than neurological. Within analytic philosophy this tradition has included the 'artificial intelligentsia' who hold that thought is the syntactic manipulation of symbols; and on the continent it has included both the structuralists and post-structuralists who insist that belief systems are properly understood in terms of the purely structural manipulation of language.

The second way to reach the sceptic's conclusion is *via* history. If the values by which we judge the veracity of our knowledge are subject to historical change then, as Nietzsche argued, how can we be sure we have the right ones? Indeed it is no longer obvious that the notion of 'being right' makes any kind of sense. The modern inheritors of this tradition are the post-modernists and radical pragmatists who insist that the truth of any claim depends solely on the values of the community in which it is made, be they scientists, religious fundamentalists, members of a

primitive tribe, or French. Moreover, once we take into account our *evolutionary*, as well as our social, history then we have the further possibility that those values should be relativised to the interests of our biological species — hence the ethical horrors of crude sociobiology or social Darwinism.

I start from the other end, from the conviction that we *can* know the world, and then ask what kind of science, both natural and social, can make sense of this ability. We do not just need a philosophy of mind that fits our science, we also need a science that fits our philosophy of mind. We must fiddle with both sides of the equation in order to get a fit. There are parallels here with Kant's transcendental approach to the problem, but I diverge from him over the question of the kinds of things we are sure we know. Kant started from the conviction that the most secure piece of knowledge we had, our exemplar or ideal of knowledge, was Newton's Laws. Kant revered Newton as Descartes revered God, and indeed they played the same role in their respective philosophies. Newton provided Kant with a rock of certainty in a swamp of doubt, just as God had done for Descartes 150 years before. But Newton, like God before him, is now clouded by doubt. There is still no denying the power of Newton's physics, but the combined effect of Lobachevsky, Einstein, Heisenberg, and Bohr, has been to show that Newton no longer provides us with the Truth *simpliciter* but, at best, an *approximation* to the truth or a *version* of the truth, or, at worst, a very useful theory. Now admitting this does not necessarily mean that we have to go all the way with the cultural relativists and claim that Newton's is just another interpretation of the world, but at the very least we have to admit that Newton's Physics is not the epistemological rock it once was.

But if God is dead and Newton injured, then who can we turn to now? Well perhaps it is time we had the confidence to start with ourselves, with our own immediate and basic convictions; and in this thesis I start from two pieces of assured knowledge that, for good or bad, should seem like 'common sense'<sup>1</sup>. The first is that we know that the things of our everyday life — rocks, trees, washing machines, other people, and so on — really exist<sup>2</sup>, and if you tell me that we do not then you must have misunderstood the words 'know' and 'exist'. But if I claim that I 'know' that rocks and washing machines and people 'exist' then the onus is on me to explain what *I* mean by 'know' and 'exist' in such a way as to avoid the sceptic's conclusions. This is the subject of the first two parts of this thesis, with the first part mostly about existing, and the second mostly about knowing.

In the first chapter I defend the view that objects are just as real as the parts that make them up. Now this may seem like an uncontroversial claim, but it runs counter to the strong reductionist current in Western materialist, and especially scientific, thought. According to this tradition, parts are in some way more basic or concrete or ontologically prior to the aggregates that they form; hence the use of the term 'fundamental' to describe the smallest particles of nature. But in chapter 2 I defend the idea of a true ontological symmetry between levels of organisation, in which rocks and trees, etc, are just as basic and fundamental and 'real' as the molecules and atoms that make them up.

The second chapter is an attempt to counter the social relativists' argument that our naive assumptions about what really exists may be very different to those from other cultures. An

---

<sup>1</sup>Though any resemblance to the philosophy of G.E.Moore starts and ends at this point.

<sup>2</sup>An attitude described by Fine (1984) as the 'natural ontological attitude'.

animist, for example, would insist that earth-spirits obviously exist, and are just as real as rocks, and so what justification do I have for claiming priority for my own view? Now I cannot prove the animist wrong, for they have their own criteria for settling such disputes, but I do have a response to the social relativist which is to make explicit the criteria by which I include rocks in my ontology and exclude spirits. Which criteria you choose is then up to you, but at least I can give a good account for my own.

In Part 2 I defend the conviction that we can know the world, that the percepts in our heads correspond to a world ‘out there’. Such correspondance theories of truth are now deeply unfashionable, and so the bulk of chapters 4–6 are spent making sense of the nature of this correspondance in a way that avoids the internalists’ various objections. I will not try to summarise the arguments here other than to mention the key starting point, which is that thought is not essentially or even primarily about internal processing — be it linguistic, computational, or neurological — but rather about our ability to physically interact with the world in meaningful ways. Internal processing is a necessary part of our ability to do this; but it is only a part.

If the first piece of sure knowledge was about facts then the second is about values. The first piece of sure knowledge was that the world exists, and the second is that we can change it for the better. In particular I am sure that I know that the values embodied in, for example, Fascism were wrong and that opposition to it was right; and if you tell me that I have no justification in fighting against oppression and exploitation and that I should accept a social order that encourages it, then you must have misunderstood the word ‘justification’. But then the problem, as before, is to find a source for these convictions, to find foundations for them. But where are we to find foundations for our ethical judgements? Kant’s strategy was to find a universal maxim, a categorical imperative, that would be beyond doubt, such that anyone who disagreed with it must be irrational. But this simply will not do. The Nazis, for example, were not irrational. The Third Reich was not the result of collective madness on the part of the German people, and nor were they the dumb dupes of Hitler’s demonic personal magnetism. On the contrary, the actions of the Nazi party were right for them and for that section of German society that they represented. Therefore reason alone cannot be the sole foundation for morality. But if we cannot appeal to universal principles to ground our value judgements then where can we turn?

In the third and fourth parts of this thesis I argue that Darwin’s theory of natural selection can supply the necessary foundations for our ethical convictions, though great care must be taken in how we apply it. The philosophical importance of Darwin’s theory is that it supplies a way of naturalising normativity, in other words it enables us to derive ethical oughts from scientific is’s. The theory of natural selection can, for example, determine what makes a heart a *good* heart. It can determine what a heart *ought* to do. According to Darwin the function of hearts is to pump blood, because if hearts had not pumped blood in the past then there would be no hearts around today. Therefore pumping blood is what hearts exist to do, it is their *purpose*. The problem then is to extend this argument from our hearts to our minds. If Darwin can determine oughts for our purely biological traits then perhaps he can do the same for our mental ones? It is at this point that we run into controversy.

The problem with extending the Darwinian argument from biological traits to mental ones is that the former are, generally, relatively fixed products of natural evolution. Our mental processes,

however, are also the products of *social* evolution; i.e. history. My solution is to argue that human history can be understood by using an approach that is *analogical* to Darwin's, namely Marx's theory of historical materialism. Now this may come as a surprise to some readers, since the application of the theory of natural selection to social processes is often associated with the political right, from the Social Darwinists to the proponents of the 'Bell Curve'. It is often forgotten that Marx himself was convinced of the deep connection his theory and Darwin's. As he wrote in a letter to Engels, 'the theory of natural selection provides the basis in natural history for our view'. The application of the theory of natural selection to social phenomenon has never been more popular, and so it seems only appropriate that we should revisit the philosopher of history who first took Darwin seriously.

However, the way in which one applies Darwin's theory to social phenomena will reflect how one applies it to biological phenomena. Therefore the whole of Part 3 is dedicated to discussing issues in evolutionary theory from a purely biological perspective, and, in particular, how we can use the theory to determine the purpose of biological traits. Then, in Part 4, I show how the argument can be extended from the biological domain to the social, and finally, hopefully, come back to the starting point which was that history does not make sure knowledge impossible. On the contrary, history is what makes it possible.

### **Acknowledgements**

I should start by thanking the giants whose shoulders I have stood on. The general idea of dialectical materialism as a way of understanding the world comes from Engels *via* Richard Lewontin and Richard Levins. The non-objectivist ontology comes from Brian Smith. The scientific realism comes from Roy Bhaskar and Elliot Sober. The physicalism comes from Richard Feynman. The evolutionary theory comes from Darwin *via* John Maynard Smith, Richard Lewontin, and Richard Dawkins. The epistemology comes from Ruth Millikan. The theory of history comes from Marx *via* Alex Callinicos and Tony Cliff. And, perhaps most importantly, the questions we should ask and the way we should answer them come from Dan Dennett. Their combined height almost compensated for my myopia.

I should also like to thank everyone in the School of Cognitive and Computing Science, the School of Biological Sciences, and the Centre for Computational Neuroscience and Robotics, for patiently answering many stupid questions; and for creating something close to the ideal of a truly interdisciplinary research atmosphere. I should also like to thank COGS for granting me the Scholarship without which this work could not have been done.

Also thanks to the Cognitive Philosophy, Artificial Life, E-Intentionality, Learning and Evolution, Evolutionary Theory, Neuroscience, and Developmental Psychology seminar groups for comments and discussion; to Ron Chrisley for trying to impose rigour; to Andy and Pepa Clark, Dan Dennett and John Maynard Smith for invaluable criticism of earlier versions; to Inman Harvey and Graham Dove for being persistently, argumentatively, and imaginatively wrong; to Nick Jakobi and Adrian Thompson for achieving in practice what I was trying to understand in theory; and to Lisa Deer for convincing me I could do it in the first place. It's her fault.