

Karl Popper

Science

and

Enlightenment

Nicholas Maxwell

Karl Popper, Science and Enlightenment

Prologue: An Idea to Help Save the World.....	4
Introduction.....	11
Chapter 1: Karl Raimund Popper: Life and Work.....	19
Chapter 2: A Critique of Popper's Views on Scientific Method..	57
Chapter 3: Science, Reason, Knowledge and Wisdom: A Critique of Specialism.....	89
Chapter 4: Popper, Kuhn, Lakatos and Aim-Oriented Empiricism.....	172
Chapter 5: The Need for a Revolution in the Philosophy of Science	234
Chapter 6: Einstein, Aim-Oriented Empiricism and the Discovery of Special and General Relativity.....	262
Chapter 7: The Enlightenment Programme and Karl Popper....	302
Chapter 8: Can the World Learn Wisdom?.....	342
Chapter 9: Non-Empirical Requirements Scientific Theories Must Satisfy: Simplicity, Unification, Explanation, Beauty..	354
Chapter 10: Scientific Metaphysics.....	375
Chapter 11: Comprehensibility rather than Beauty.....	410
Chapter 12: A Mug's Game? Solving the Problem of Induction with Metaphysical Presuppositions.....	437
Chapter 13: Does Probabilism Solve the Great Quantum Mystery?.....	474
References.....	496

Acknowledgements

The Prologue: ‘An Idea to Help Save the World’, adapted from the Preface to the 2nd edition of *What’s Wrong With Science?*, was first published in *Sublime*, Issue 17, 2009, pp. 90-93.

Chapter I: ‘Karl Raimund Popper: Life and Work’ was first published, in a slightly modified form, in *British Philosophers, 1800-2000*, ed., P. Dematteis, P. Fosl and L. McHenry, Brucoli Clark Layman, Columbia, 2002, pp. 176-194.

Chapter 2; ‘A Critique of Popper’s Views on Scientific Method’ was first published in *Philosophy of Science* 39, 1972, pp. 131-52.

Chapter 3: ‘Science, Reason, Knowledge and Wisdom: A Critique of Specialism’ was first published in *Inquiry* 23, 1980, pp. 19-81.

Chapter 4: ‘Popper, Kuhn, Lakatos and Aim-Oriented Empiricism’ was first published in *Philosophia* 32, nos. 1-4, 2005, pp. 181-239.

Chapter 5: ‘The Need for a Revolution in the Philosophy of Science’ was first published in *Journal for General Philosophy of Science* 33, 2002, pp. 381-408.

Chapter 6: ‘Einstein, Aim-Oriented Empiricism and the Discovery of Special and General Relativity’ was first published in *The British Journal for the Philosophy of Science* 44, 1993, pp. 275-305.

Chapter 7: ‘The Enlightenment Programme and Karl Popper’ was first published, in a shortened form, in *Karl Popper: A Centenary Assessment. Volume 1: Life and Times, Values in a World of Facts*, ed. I. Jarview, K. Milford and D. Miller, chapter 11, Ashgate, London, pp. 177-190.

Chapter 8: ‘Can the World Learn Wisdom’ was first published in *Solidarity, Sustainability, and Non-Violence*, vol. 3, no 4, April 2007, <http://www.pelicanweb.org/solisustv03n04maxwell.html>

Chapter 13: ‘Does Probabilism Solve the Great Quantum Mystery?’, was first published in *Theoria* vol. 19/3, no. 51, pp. 321-336.

I am grateful to the publishers and editors of the above books and journals for permission to reprint these essays.

Chapters 9-12 have not been published before.

Prologue

An Idea to Help Save the World

(Adapted from the Preface to the 2nd edition of *What's Wrong With Science?*, and first published in *Sublime*, Issue 17, 2009, pp. 90-93.)

Here is an idea that just might save the world. It is that science, properly understood, provides us with the methodological key to the salvation of humanity.

A version of this idea can be found buried in the works of Karl Popper. Famously, Popper argued that science cannot verify theories, but can only refute them. This sounds very negative, but actually it is not, for science succeeds in making such astonishing progress by subjecting its theories to sustained, ferocious attempted falsification. Every time a scientific theory is refuted by experiment or observation, scientists are forced to try to think up something better, and it is this, according to Popper, which drives science forward.

Popper went on to generalize this falsificationist conception of scientific method to form a notion of rationality, *critical rationalism*, applicable to all aspects of human life. Falsification becomes the more general idea of *criticism*. Just as scientists make progress by subjecting their theories to sustained attempted empirical falsification, so too all of us, whatever we may be doing, can best hope to achieve progress by subjecting relevant ideas to sustained, severe *criticism*. By subjecting our attempts at solving our problems to criticism, we give ourselves the best hope of discovering (when relevant) that our attempted solutions are inadequate or fail, and we are thus compelled to try to think up something better. By means of judicious use of criticism, in personal, social and political life, we may be able to achieve, in life, progressive success somewhat like the progressive success achieved by science. We can, in this way, in short, learn from

scientific progress how to make personal and social progress in life. Science, as I have said, provides the methodological key to our salvation.

I discovered Karl Popper's work when I was a graduate student doing philosophy at Manchester University, in the early 1960s. As an undergraduate, I was appalled at the triviality, the sterility, of so-called "Oxford philosophy". This turned its back on all the immense and agonizing problems of the real world – the mysteries and grandeur of the universe, the wonder of our life on earth, the dreadful toll of human suffering – and instead busied itself with the trite activity of analysing the meaning of words. Then I discovered Popper, and breathed a sigh of relief. Here was a philosopher who, with exemplary intellectual integrity and passion, concerned himself with the profound problems of human existence, and had extraordinarily original and fruitful things to say about them. The problems that had tormented me had in essence, I felt, already been solved.

But then it dawned on me that Popper had failed to solve his fundamental problem – the problem of understanding how science makes progress. In one respect, Popper's conception of science is highly unorthodox: all scientific knowledge is conjectural; theories are falsified but cannot be verified. But in other respects, Popper's conception of science is highly orthodox. For Popper, as for most scientists and philosophers, the basic aim of science is knowledge of truth, the basic method being to assess theories with respect to evidence, *nothing being accepted as a part of scientific knowledge independently of evidence*. This orthodox view – which I came to call *standard empiricism* – is, I realised, *false*. Physicists only ever accept theories that are *unified* – theories that depict the same laws applying to the range of phenomena to which the theory applies. Endlessly many empirically more successful *disunified* rivals can always be concocted, but these are always ignored. This means, I realised, that science does make a big, permanent, and highly problematic assumption about the nature of the universe independently of empirical considerations and even, in a sense, in violation of empirical considerations – namely, that the universe is such that all grossly *disunified* theories are false. Without some

such presupposition as this, the whole empirical method of science breaks down.

It occurred to me that Popper, along with most scientists and philosophers, had misidentified the basic aim of science. This is not truth *per se*. It is rather truth *presupposed to be unified*, presupposed to be explanatory or comprehensible (unified theories being *explanatory*). Inherent in the aim of science there is the metaphysical – that is, untestable – assumption that there is some kind of underlying *unity* in nature. The universe is, in some way, physically comprehensible.

But this assumption is profoundly problematic. We do not *know* that the universe is comprehensible. This is a conjecture. Even if it is comprehensible, almost certainly it is not comprehensible in the way science presupposes it is today. For good Popperian reasons, this metaphysical assumption must be made explicit within science and subjected to sustained *criticism*, as an integral part of science, in an attempt to improve it.

The outcome is a new conception of science, and a new kind of science, which I called *aim-oriented empiricism*. This subjects the aims, and associated methods, of science to sustained critical scrutiny, the aims and methods of science evolving with evolving knowledge. Philosophy of science (the study of the aims and methods of science) becomes an integral, vital part of science itself. And science becomes much more like natural philosophy in the time of Newton, a synthesis of science, methodology, epistemology, metaphysics and philosophy.

The aim of seeking *explanatory truth* is however a special case of a more general aim, that of seeking *valuable truth*. And this is sought in order that it be *used* by people to enrich their lives. In other words, in addition to metaphysical assumptions inherent in the aims of science there are *value* assumptions, and *political* assumptions, assumptions about how science should be used in life. These are, if anything, even more problematic than metaphysical assumptions. Here, too, assumptions need to be made explicit and critically assessed, as an integral part of science, in an attempt to improve them.

Released from the crippling constraints of standard empiricism, science would burst out into a wonderful new life, realising its full

potential, responding fully both to our sense of wonder and to human suffering, becoming both more rigorous and of greater human value.

And then, in a flash of inspiration, I had my great idea. I could tread a path parallel to Popper's. Just as Popper had generalized falsificationism to form critical rationalism, so I could generalise my aim-oriented empiricist conception of scientific method to form an aim-oriented conception of rationality, potentially fruitfully applicable to all that we do, to all spheres of human life. But the great difference would be this. I would be starting out from a conception of science – of scientific method – that enormously improves on Popper's notion. In generalizing this, to form a general idea of progress-achieving rationality, I would be creating an idea of immense power and fruitfulness.

I knew already that the line of argument developed by Popper, from falsificationism to critical rationalism, was of profound importance for our whole culture and social order, and had far-reaching implications and application for science, art and art criticism, literature, music, academic inquiry quite generally, politics, law, morality, economics, psychoanalytic theory, evolution, education, history – for almost all aspects of human life and culture. The analogous line of argument I was developing, from aim-oriented empiricism to aim-oriented rationalism, would have even more fruitful implications and applications for all these fields, starting as it did from a much improved initial conception of the progress-achieving methods of science.

The key point is extremely simple. It is not just in science that aims are profoundly problematic. This is true in life as well. Above all, it is true of the aim of creating a good world – an aim inherently problematic for all sorts of more or less obvious reasons. It is not just in science that problematic aims are misconstrued or “repressed”; this happens all too often in life too, both at the level of individuals, and at the institutional or social level as well. We urgently need to build into our scientific institutions and activities the aims-and-methods-improving methods of aim-oriented empiricism, so that scientific aims and methods improve as our scientific knowledge and understanding improve. Likewise, and even more urgently, we need to build into

all our other institutions, into the fabric of our personal and social lives, the aims-and-methods-improving methods of aim-oriented rationality, so that we may improve our personal, social and global aims and methods as we live.

One outcome of the 20th century is a widespread and deep-seated cynicism concerning the capacity of humanity to make real progress towards a genuinely civilized, good world. Utopian ideals and programmes, whether of the far left or right, that have promised heaven on earth, have led to horrors. Stalin's and Hitler's grandiose plans led to the murder of millions. Even saner, more modest, more humane and rational political programmes, based on democratic socialism, liberalism, or free markets and capitalism, seem to have failed us. Thanks largely to modern science and technology, many of us today enjoy far richer, healthier and longer lives than our grandparents or great grandparents, or those who came before. Nevertheless the modern world is confronted by grave global problems: the lethal character of modern war, the spread and threat of armaments, conventional, chemical, biological and nuclear, rapid population growth, severe poverty of millions in Africa, Asia and elsewhere, destruction of tropical rain forests and other natural habitats, rapid extinction of species, annihilation of languages and cultures. And over everything hangs the menace of climate change, threatening to intensify all the other problems (apart, perhaps, from population growth).

All these grave global problems are the almost inevitable outcome of the successful exploitation of science and technology plus the failure to build aim-oriented rationality into the fabric of our personal, social and institutional lives. Modern science and technology make modern industry and agriculture possible, which in turn make possible population growth, modern armaments and war, destruction of natural habitats and extinction of species, and global warming. Modern science and technology, in other words, make it possible for us to achieve the goals of more people, more industry and agriculture, more wealth, longer lives, more development, housing and roads, more travel, more cars and aeroplanes, more energy production and use, more and more lethal armaments (for defence only of course!). These things seem

inherently desirable and, in many ways, are highly desirable. But our successes in achieving these ends also bring about global warming, war, vast inequalities across the globe, destruction of habitats and extinction of species. All our current global problems are the almost inevitable outcome of our long-term failure to put aim-oriented rationality into practice in life, so that we actively seek to discover problems associated with our long-term aims, actively explore ways in which problematic aims can be modified in less problematic directions, and at the same time develop the social, the political, economic and industrial *muscle* able to change what we do, how we live, so that our aims become less problematic, less destructive in both the short and long term. We have failed even to appreciate the fundamental need to improve aims and methods as the decades go by. Conventional ideas about rationality are all about *means*, not about *ends*, and are not designed to help us *improve* our ends as we proceed. Implementing aim-oriented rationality is essential if we are to survive in the long term. To repeat, the idea spelled out in this book, if taken seriously, just might save the world.

Einstein put his finger on what is wrong when he said "Perfection of means and confusion of goals seems, to my opinion, to characterize our age." This outcome is inevitable if we restrict rationality to *means*, and fail to demand that rationality – the authentic article – must quite essentially include the sustained critical scrutiny of *ends*.

Scientists, and academics more generally, have a heavy burden of responsibility for allowing our present impending state of crisis to develop. Putting aim-oriented rationality into practice in life can be painful, difficult and counter-intuitive. It involves calling into question some of our most cherished aspirations and ideals. We have to *learn* how to live in aim-oriented rationalistic ways. And here, academic inquiry ought to have taken a lead. The primary task of our schools and universities, indeed, ought to have been, over the decades, to help us learn how to improve aims and methods as we live. Not only has academia failed miserably to take up this task, or even see it as necessary or desirable. Even worse, perhaps, academia has failed itself to put aim-oriented rationality into practice. Science has met with such astonishing

success because it has put something like aim-oriented empiricism into scientific practice – but this has been obscured and obstructed by the conviction of scientists that science ought to proceed in accordance with standard empiricism – with its fixed aim and fixed methods. Science has achieved success despite, and not because of, general allegiance of scientists to standard empiricism.

The pursuit of scientific knowledge dissociated from a more fundamental concern to help humanity improve aims and methods in life is, as we have seen, a recipe for disaster. This is the crisis behind all the others. We are in deep trouble. We can no longer afford to blunder blindly on our way. We must strive to peer into the future and steer a course less doomed to disaster. Humanity must learn to take intelligent and humane responsibility for the unfolding of history.

Introduction

Karl Popper is famous for having proposed that science advances by a process of conjecture and refutation. He is also famous for defending the open society against what he saw as its arch enemies – Plato and Marx.

Popper's contributions to thought are of profound importance, but they are not the last word on the subject. They need to be improved. My concern in this book is to spell out what is of greatest importance in Popper's work, what its failings are, how it needs to be improved to overcome these failings, and what implications emerge as a result.

The basic theme of the book has already been summarized in the Prologue. In what follows I spell out this theme in greater detail. The book consists of a collection of essays which dramatically develop Karl Popper's views about natural and social science, and how we should go about trying to solve social problems.

Criticism of Popper's falsificationist philosophy of natural science leads to a new philosophy of science, which I call aim-oriented empiricism.¹ This makes explicit metaphysical theses concerning the comprehensibility and knowability of the universe that are an implicit part of scientific knowledge – implicit in the way science excludes all theories that are not explanatory, even those that are more successful empirically than accepted theories. Aim-oriented empiricism has major implications, not just for the academic discipline of philosophy of science, but for science itself.

Popper generalized his philosophy of science of falsificationism to arrive at a new conception of rationality – critical rationalism – the key methodological idea of Popper's profound critical exploration of political and social issues in his *The Open Society and Its Enemies*, and *The Poverty of Historicism*. This path of Popper, from scientific method to rationality and social and political issues is followed here, but the starting point is aim-

¹ Aim-oriented empiricism was first put forward in my (1974). Anyone interested in the way the view has evolved over the years should have a look at this two-part paper.

oriented empiricism rather than falsificationism. Aim-oriented empiricism is generalized to form a new conception of rationality – aim-oriented rationalism – which has far-reaching implications for political and social issues, for the nature of social inquiry and the humanities, and indeed for academic inquiry as a whole. The strategies for tackling social problems that arise from aim-oriented rationalism improve on Popper's recommended strategies of piecemeal social engineering and critical rationalism, associated with Popper's conception of the open society. This book thus sets out to develop Popper's philosophy in new and fruitful directions.

The theme of the book, in short, is to discover what can be learned from scientific progress about how to achieve social progress towards a better world. That there is indeed much to be learned from scientific progress about how to achieve social progress was the big idea of the 18th century Enlightenment. This was immensely influential. But the *philosophes* of the Enlightenment made mistakes, and these mistakes, inherited from the Enlightenment, are built into the institutional and intellectual structure of academic inquiry today. In his two great works, *The Logic of Scientific Discovery* and *The Open Society and Its Enemies*, Popper corrected some of the mistakes of the Enlightenment – mistakes about the nature of scientific method and rationality. But Popper left other mistakes undetected and uncorrected. The present book seeks to push the Popperian research programme further, and correct what Popper left uncorrected.

The fundamental idea that emerges is that there is an urgent need to bring about a revolution in academic inquiry so that it takes up its proper task of promoting *wisdom* and not just acquiring knowledge – wisdom being the capacity to realize what is of value in life for oneself and others, thus including knowledge and technological know-how, but much else besides. I have devoted much of my working life to trying to get this idea across. The essays that follow provide a record of this life work.

Most philosophers of science see their work as contributing to a *meta-discipline*. The object of study is science, and the task is to describe, explain and understand this object. It is no more the proper task of the philosopher of science to *criticize* science, or to

make suggestion as to how science can be *improved*, than it is the task of the astronomer to criticize the moon. But this standard *meta-discipline* way of conceiving the subject entirely misconstrues what ought to be the proper relationship between science and the philosophy of science. A major implication of the view to be expounded and defended here, aim-oriented empiricism, is that the rationality of science requires that the philosophy of science – critical exploration of views concerning the aims and methods of science – is an integral, influential part of science itself, both being influenced by, and influencing, science. In other words, in order to be rigorous, science must include some imaginative and critical exploration of problematic aims and methods. The very act of setting up the philosophy of science as a *meta-discipline*, distinct from science itself, looking down on science from above, as it were, describing and seeking to explain and understand what goes on, but in no way interfering with, contributing to or criticizing science, serves to undermine the very thing the discipline seeks to understand, namely the rationality of science. The orthodox *meta-discipline* approach not only makes the subject sterile (in that it can have nothing to contribute to science itself); it makes it quite impossible to solve the fundamental problem of the discipline – the rationality of science. Indeed, the discipline, so conceived, actually becomes the source of a pervasive and damaging *irrationality* in science.

I make no apology, therefore, for criticizing science, for attempting to contribute to and improve science, in what follows. Philosophy of science pursued within the framework of aim-oriented empiricism might be compared to the work Weierstrass, Dedekind and others in the late 19th and early 20th century in bringing greater rigour to mathematics: they made mathematics more rigorous, and contributed to mathematics itself, at one and the same time. Somewhat analogously, I seek to increase the rigour of science, and make a contribution to science at the same time. I might add that in criticizing science and suggesting how it might be improved (made more rigorous and of greater human value) I am again developing a tendency to be found scattered among Popper's works. Despite – or perhaps because of – his great admiration for science at its best, Popper does not shrink

from criticizing what he sees as deplorable aspects of science: specialization, authoritarianism, submission to mere intellectual fashion, failure to grapple with the fundamental problems of cosmology. Again, Popper depicts graphically some of the bad consequences, for science itself, of attempting to put bad inductivist methods into scientific practice. But all this is paradoxically at odds with a major tenet of Popper's philosophy of science, namely his proposed solution to the problem of demarcation. Popper holds that an idea, in order to be scientific, must be empirically *falsifiable*. Philosophies of science, because they are not empirically falsifiable, are not a part of science. They are to be severely *demarcated* from science. Thus Popper, in actively criticizing aspects of science, violates the precepts of his own philosophy of science. All this changes dramatically once Popper's philosophy of science has been amended to become the doctrine espoused here, aim-oriented empiricism.²

The points just made concerning the proper relationship between science on the one hand and the philosophy of science on the other, will turn out to have a major bearing on further developments of the argument concerning social inquiry. I argue that social inquiry needs to be construed, not primarily as social *science*, but rather as social *philosophy* or social *methodology*. Social inquiry is not to be related to the social world as astronomy is to the moon, or geology is to the earth. Social inquiry is not, fundamentally, engaged in seeking to acquire knowledge about social phenomena. Rather, social inquiry needs to take the relationship between science and the philosophy of science (as specified by aim-oriented empiricism) to be the model, the ideal, of how social inquiry ought to be related to society. What the philosophy of science is to science, so social inquiry is to society. The proper task of social inquiry is to help worthwhile social endeavours improve their problematic aims and methods as they proceed, just as the proper task of the philosophy of science is to help science improve its

² For a discussion of these contradictory impulses to be found in Popper's work, together with relevant references to Popper's writings, see my (2011).

problematic aims and methods as it proceeds. On this view, indeed, the philosophy of natural science is just that small, but crucial bit of social inquiry that deals with the worthwhile social endeavour of *natural science*.

Let me now indicate, in a little more detail, the contents of the chapters of this book.

Chapter One gives an account of Karl Popper's life and work. I make it quite clear that, in my view, Popper is the greatest philosopher of the 20th century. I am nevertheless critical of aspects of his work – it would be a betrayal of his “critical philosophy” not to be. In a preliminary way, I indicate what are, in my view, unsolved problems inherent in his views, and what needs to be done to overcome these difficulties. Subsequent chapters seek to improve Popper's philosophy in some key respects in order to overcome these defects.

Chapter Two criticizes Popper's falsificationist conception of scientific method. I argue that criticisms of Thomas Kuhn, Imre Lakatos, Paul Feyerabend and others are not too serious, and can readily be answered. A more serious criticism is that Popper has failed to provide a rationale for the methods he advocates.

Chapter Three criticizes specialism, the doctrine that academic inquiry quite properly consists of a great number of specialized disciplines, only specialized intellectual standards being worthwhile. This is opposed by the view that academic inquiry must engage in sustained exploration of fundamental problems that cut across disciplinary boundaries, this exploration being undertaken in a way which influences, and is influenced by, specialized research. This is an improvement of anti-specialist remarks scattered throughout Popper's works.

Chapter Four expounds and argues for aim-oriented empiricism, a major improvement over Popper's falsificationist conception of science. Theoretical physics only ever accepts unified theories, even though endlessly many empirically more successful, disunified rivals can always be formulated. This means, I argue, that fundamental theoretical physics (and therefore all of natural science) implicitly accepts that there is some kind of underlying unity in the physical universe. It is, in some way or other, physically comprehensible. But this untestable (or metaphysical)

assumption, inherent in the aims of science, precisely because it is problematic, influential and implicit, needs to be made explicit within the context of science, so that it can be critically assessed, so that alternatives can be developed and considered, in an attempt to improve that version of the assumption that is adopted. In order to facilitate improvement of this highly problematic assumption, we need, I argue, to represent it in the form of a hierarchy of assumptions and associated methods, assumptions becoming progressively less and less substantial, and more and more such that their truth is required for science, or the pursuit of knowledge, to be possible at all. It is this hierarchical methodology that is the nub of aim-oriented empiricism. I argue that aim-oriented empiricism solves the problem of induction – the problem of the rationality of science – the outstanding fundamental problem in the philosophy of science, and provides science with a rational, if fallible, method of discovery. The hierarchical methodology of aim-oriented empiricism provides science with something like positive feedback between knowledge and methods, so that as scientific knowledge and understanding improve, the aims and methods of science improve as well – knowledge about how to improve knowledge, in other words. It is this, according to aim-oriented empiricism, that is a key element of scientific rationality, and helps account for the explosive growth in scientific knowledge since modern science began in the 17th century.

In chapter Four I argue that aim-oriented empiricism is a kind of synthesis of the views of Popper, Kuhn and Lakatos, but also a dramatic improvement over all three views.

Aim-oriented empiricism stands in sharp contrast to standard empiricism, one version or another of which is taken for granted by most scientists and philosophers of science. According to standard empiricism, the basic intellectual aim of science is truth, and the basic method is to assess claims to knowledge of truth impartially with respect to evidence. Considerations of simplicity, unity or explanatory character may legitimately influence preference for a theory for a time, but not in such a way that the universe itself is presumed to be simple, unified or comprehensible. The key tenet of all versions of standard empiricism is that no assumption about the universe can be accepted as a part of scientific knowledge

independent of evidence, let alone in conflict with evidence. This stands in stark contrast with aim-oriented empiricism, which holds that persistent scientific acceptance of unified theories means that science implicitly accepts that the universe itself possesses some kind of underlying unity.

Chapter Five argues for the urgent need for scientists and philosophers of science alike to reject standard empiricism and adopt aim-oriented empiricism instead, and provides an account of the revolution in our whole conception of science that this would involve.

Chapter Six argues that Einstein was the first scientist to put something like aim-oriented empiricism explicitly into scientific practice in discovering special and general relativity. The method of discovery of aim-oriented empiricism played a crucial role in Einstein's discovery of these theories. And not only did Einstein implement aim-oriented empiricism in scientific practice: after his discovery of general relativity, Einstein came to *advocate* a view that came closer and closer to aim-oriented empiricism, as the years passed.

Chapter Seven argues that Popper's *The Logic of Scientific Discovery* and *The Open Society and Its Enemies*, taken together, constitute a major development of the Enlightenment programme of learning from scientific progress how to achieve social progress towards a more enlightened world. But what Popper has to say is not the last word on the subject. Popper's version of the Enlightenment programme needs further improvement, partly because Popper's conception of scientific method needs to be improved, but mainly because, in order to implement the programme, we need to apply scientific method, not to social science, but to *the social world itself*. How and why Popper's version of the Enlightenment programme needs to be improved is outlined in this chapter.

Chapter Eight argues that it is essential to put aim-oriented rationality (generalized from aim-oriented empiricism) into academic practice if humanity is to have what it so urgently needs, a kind of inquiry rationally designed to help it learn how to become more civilized.

Chapter Nine solves the problem of what it means to say of a scientific theory that it is simple, unified or explanatory. This problem was recognized by Popper (see his *Conjectures and Refutations*, p. 241), but Popper did not know how to solve it. Einstein recognized the problem too, but did not know how to solve it either. It is one of the great successes of aim-oriented empiricism that it provides the means for the problem to be solved.

Chapter Ten restates the argument for aim-oriented empiricism, and the solution it provides to the problem of simplicity of theory, and shows how this view provides the means for partially ordering theories and metaphysical theses in terms of their degree of unity.

Chapter Eleven compares and contrasts views about simplicity, unity, explanatory power or “beauty” associated with aim-oriented empiricism on the one hand, and a view put forward by James McAllister on the other hand (see his influential *Beauty and Revolution in Science*).

Chapter Twelve argues that aim-oriented empiricism succeeds in doing what Popper’s falsificationism fails to do, namely solve the problem of induction.

Chapter Thirteen takes up a theme close to Popper’s heart: the problems of interpreting quantum theory in a realist way – so that the theory can be understood to be about electrons, nuclei, atoms and other denizens of the quantum world, and is not doomed to be a mere *instrument* for the prediction of experimental results. I argue that *probabilism* is the key to solving the fundamental quantum mystery – the apparent capacity of quantum entities (electrons, atoms and so on) to behave *both* as particles and waves. Probabilism, here, is the doctrine that nature herself is probabilistic. What exists at one moment may only determine what exists next *probabilistically*, and not deterministically. This develops basic ideas of Popper about quantum theory, but in ways of which he strongly disapproved.

Chapter One

Karl Raimund Popper

(Draft of contribution to *British Philosophers, 1800-2000*, edited by P. Dematteis, P. Fosl and L. McHenry, Brucoli Clark Layman, Columbia, 2002, pp. 176-194.)

Karl Popper is the greatest philosopher of the 20th century. No other philosopher of the period has produced a body of work that is as significant. What is best in Popper's output is contained in his first four published books. These tackle fundamental problems with ferocious, exemplary integrity, clarity, simplicity and originality. They have widespread, fruitful implications, for science, for philosophy, for the social sciences, for education, for art, for politics and political philosophy.

In his first published book *The Logic of Scientific Discovery* (1959, first published in German in 1934), Popper argues that, although scientific theories cannot be verified, or even rendered probable, by evidence, they can be falsified. Science makes progress by putting forward falsifiable conjectures – theories which say as much as possible about the world, and which thus expose themselves as much as possible to the risk of empirical refutation; they are then subjected to a ruthless onslaught of attempted observational and experimental refutation. When finally a scientific theory is falsified empirically, the task then becomes to think up an even better theory, which says even more about the world. The new theory must predict all the success of the old theory, predict successfully the phenomena that falsified the old theory, and predict new phenomena as well. In his next book, *The Open Society and Its Enemies* (1966), written during the second war and first published in 1945, Popper tackles problems that arise in connection with creating an "open" society, one which tolerates diversity of views and ways of life. Popper argues that some of the greatest thinkers have been opposed to the "open" society, most

notably Plato and Marx. In *The Poverty of Historicism* (1961), first published in 1957, Popper is concerned to demolish the view that social science should, or can, predict the way societies evolve. Popper spells out his view of how social science should be developed, closely modelled on the account of natural science given in *The Logic of Scientific Discovery*. The next book, *Conjectures and Refutations* (1963), is a collection of essays which restate in a more accessible way Popper's falsificationist view of science, and draw out implications for a range of philosophical problems. Further books include *Objective Knowledge* (1972), a collection of essays which draw on the analogy between Darwinian evolution and scientific progress, and which expound Popper's view that there exists, in addition to the material world and the psychological world, a third world of theories, problems and arguments; *The Self and Its Brain* (1977), written with the neurologist John Eccles, which applies Popper's "third world" view to the mind-body problem; and the three volume *The Postscript* (1982, 1982a, 1983), which amounts to a massive restatement and development of Popper's falsificationist conception of science. A volume of *The Library of Living Philosophers* (Schilpp, 1974) is devoted to Popper's work; this includes Popper's intellectual autobiography, published subsequently as a separate book with the title *Unended Quest* (1976). This gives a fascinating and gripping account of the development of Popper's thought.

Fundamental to Popper's philosophy is the idea that criticism lies at the heart of rationality. It would be a betrayal of Popper's philosophy to give an entirely uncritical exposition of his work; some criticism of key tenets of his philosophy will therefore be included in what follows.

Karl Raimund Popper was born in Vienna on 28th July 1902. His parents were Jewish but converted to Protestantism before their children were born. Popper's father, Simon Carl Siegmund (1856-1932), was a doctor of law of the University of Vienna. He had a successful legal practice in Vienna, at which he apparently worked hard, but his real interests lay in the direction of scholarship and literature. Popper's mother, Jenny Schiff (1864-1938), came from a musical family, and was herself musical.

Popper tells us that she played the piano beautifully; music had an important place in Popper's life.

During Popper's early childhood, his parents were prosperous. They lived in a large apartment in an 18th century house in the centre of Vienna, where Popper's father conducted his legal practice. Popper's father had an enormous library, which included many works of philosophy; books were everywhere, Popper tells us, except in the dining room, where stood a concert grand piano.

As a young boy, Popper was much concerned with the poverty he saw all around him in Vienna. In his autobiography, Popper recounts an early brush with philosophy. His father had suggested he read some volumes of Strinberg's autobiography. Finding that Strinberg gave much too much importance to words and their meanings, Popper tried to point this out to his father, and was surprised to discover that he did not agree. Popper, later, saw this as his first brush with a life-long battle to combat the influential view that philosophy must concern itself with analysis of meaning.

Popper left school at 16 because of the tedium of the classes, and enrolled at the University of Vienna, initially as a non-Matriculated student. Four years later, at the second attempt, he passed the exam to become a Matriculated student. Any student could take any lecture course. Initially, Popper sampled lectures in a wide range of subjects – history, literature, psychology, philosophy – but then concentrated on physics and mathematics. In these fields Popper had excellent, if remote and autocratic, teachers: Hans Thirring, Wirtinger, Furtwängler and Hans Hahn. Later, Popper devoted himself to the study of the psychology of thinking, influenced by Karl Bühler and the writings of Otto Selz.

The first world war and its aftermath brought dramatic changes to conditions of life in Vienna. Popper's father lost much of his savings. Popper left home and moved into part of a disused military hospital converted by students into a primitive students' home, and joined socialist groups seeking political change. For a time, Popper thought of himself as a communist. But an event then occurred which Popper was later to describe as one of the most important in his life, which caused him to become critical of communism and Marxism, and which, years later, led to the writing of his *The Open Society and Its Enemies*. The communists

organized a demonstration with the intention of freeing communists held in a police station in Vienna. The police opened fire, and some of the demonstrators were killed. Popper was deeply shocked, and even felt some personal responsibility for the tragedy, in that he had endorsed a doctrine, Marxism, which required that there should be just such incidents, so that the struggle to overcome capitalism might be intensified.

Popper nevertheless continued to think of himself as a socialist, and to associate with socialist groups. In his autobiography, Popper celebrates these groups of working people for their dedication, their eagerness to become educated. Even though the times were troubled, the economic and political outlook bleak, Popper says that he and his friends were often exhilarated at the intellectual and political challenges that lay before them. For a time Popper worked as a labourer, but found the work too hard; he then tried his hand at cabinet making, but was distracted by the intellectual problems that he was working on. Popper also worked for the psychologist Adler, and as a social worker concerned with neglected children.

Shortly before submitting his dissertation for his Ph.D., the focus of Popper's interest switched from the psychology, to the methodology, of thought and problem-solving, and in particular to the methodology of science. This came about partly as a result of long discussions with two friends, the philosophers Julius Kraft and Heinrich Gomperz. The dissertation was hastily written; Popper's examiners were Bühler and Schlick; Popper thought he had failed, but in fact he passed with distinction.

At this time a Pedagogic Institute was created in Vienna to train teachers in new methods of education. Popper decided to become a teacher, joined the course, held informal seminars for fellow students, and duly became qualified to teach physics and mathematics in secondary schools. He met, and later married, a fellow student, Josephine Henninger (Hennie), who also became a teacher.

While employed full-time as a teacher, Popper continued to work hard at epistemological and methodological problems of science, writing down his thoughts as an aid to research, rather than with the idea that the work might eventually be published.

During this time, Popper got to know a number of people associated with the Vienna Circle, famous for promoting logical positivism. The Vienna Circle was essentially a seminar which one attended when invited by its convenor, Moritz Schlick. Rudolf Carnap, Otto Neurath, Herbert Feigl, Kurt Gödel, Friedrich Waismann, Victor Kraft, Karl Menger, Hans Hahn, Philipp Frank, Richard von Mises, Hans Reichenbach and Carl Hempel were among the members; Ludwig Wittgenstein, much admired by Schlick, was the intellectual godfather (together, perhaps, with Ernst Mach and Bertrand Russell). Visitors from abroad included A. J. Ayer and Frank Ramsey from England, Ernst Nagel and W. V. Quine from the USA, Arne Naess from Norway, and Alfred Tarski from Poland. But Popper was never invited to join the Circle (possibly because Schlick was aware of Popper's low opinion of Wittgenstein). Nevertheless, Popper did attend, and give papers at, a number of fringe seminars, and Popper's work was strongly influenced by, but also critical of, the doctrines of the Circle.

Two issues were of central concern to Popper. The first was the problem of how to distinguish science from pseudo-science. Popper was impressed by the difference between the theories of Marx, Freud and Adler on the one hand, and Einstein's general theory of relativity, on the other. The former theories seemed able to explain phenomena whatever happened; nothing, it seemed, could tell against these theories. Einstein's theory, by contrast, issued in a definite prediction; light travelling near the sun would pursue a curved path due to the gravitational field of the sun. If this did not happen, Einstein's theory would be refuted. Popper decided, around 1921 (he tells us) that this constituted the key difference between pseudo and genuine scientific theories: whereas the former were unrefutable, the latter were open to empirical refutation.

The other problem that preoccupied Popper was that of the logic, or methodology, of scientific discovery: How does science acquire new knowledge? This was the problem that confronted Popper when his earlier interest in the psychology of thinking transmuted into interest in the logic of thinking, the logic of discovery.

Suddenly, Popper tells us, he put two and two together. His earlier solution to the first problem also solves the second problem. There is no such thing as the verification of theories in science; there is only refutation. Scientists put forward theories as empirically falsifiable conjectures or guesses: these are then subjected to sustained attempted empirical refutation. Science advances through a process of trial and error, of conjecture and refutation.

Encouraged by Feigl, Popper wrote the first volume of what was intended to be a two volume work, entitled *Die beiden Grundprobleme der Erkenntnistheorie* (*The Two Fundamental Problems of the Theory of Knowledge*). The first volume was accepted for publication by Schlick and Frank, the editors of a series of publications written mostly by members of the Vienna Circle. The publisher, Springer, insisted the book must be shortened. But in the meantime Popper had finished the second volume. He offered a new work consisting of extracts from both volumes; but this was still judged by Springer to be too long. Popper's uncle, Walter Schiff, cut the manuscript by about a half, and this was finally published late in 1934 as *Logic der Forschung*. Thus emerged into the public domain, in the shadow of Hitler and impending war, what is, perhaps, the most important book on scientific method to be published in the last century. The book was only published in English translation, as *The Logic of Scientific Discovery* (with many additional appendices and footnotes), in 1959. (*Die beiden Grundprobleme der Erkenntnistheorie* was not published until much later, in 1979.)

The Logic of Scientific Discovery (*L.Sc.D.*) begins by spelling out what are, for Popper, the two fundamental problems concerning the nature of scientific inquiry. (1) The problem of induction: How can scientific theories be verified by evidence, in view of Hume's arguments which seem to show that this is impossible? (2) The problem of demarcation: How is science to be demarcated from non-science (pseudo science and metaphysics)? As we saw above, Popper's solution to the second problem is that, in order to be scientific, a theory must be empirically *falsifiable*. This, for Popper, solves the first problem as well. Scientific laws and theories cannot be verified by evidence at all; they can only be

falsified. However much evidence may be amassed in support of a theory, its probability remains zero. But despite this negative conclusion, science can still make progress. This comes about as a result of theories being proposed as conjectures, in response to problems: these conjectures are then subjected to a ruthless barrage of attempted empirical refutation. The purpose of observation and experimentation is not to verify, but to refute. When a theory is refuted empirically, this creates the problem of discovering a new conjecture, a new theory, even more successful than its predecessor in that it meets with all the success of its predecessor, successfully predicts the phenomena that refuted its predecessor, and predicts new phenomena as well. When such a theory is formulated, the task then becomes to try to refute this new theory in turn. Thus science advances, from one falsifiable conjecture to another, each successfully predicting more than its predecessor, but none ever having probability greater than zero. All theoretical knowledge in science is irredeemably conjectural in character. But science makes progress precisely because, in science, it is possible to discover that theories are false, and thus need to be replaced by something better.

Popper has been much criticized for not appreciating that even empirical refutations are not decisive: it is always a conjecture that a theory has been falsified, since it is always a conjecture that a given observation or experiment has yielded a falsifying result. But Popper has at least two replies to such criticisms.

First, there is a decisive logical asymmetry between verification and falsification. Any theory has infinitely many empirical consequences, for infinitely many times and places. We, however, can only ever verify finitely many of these consequences, and thus must forever be infinitely far away from verifying the theory. But we only need to discover one false empirical consequence of a theory in order to show decisively that the theory is false.

Second, Popper emphasizes that a theory is only falsifiable with respect to the adoption of a methodology. Given that a theory is empirically falsified, it is always possible to rescue the theory from falsification by adopting what Popper calls "conventionalist stratagems". These include explaining the experimental result away in some way, or modifying the theory, in an *ad hoc* way, so

that it no longer clashes with the empirical result. Popper proposes that science should adopt methodological rules governing the way theories are to be accepted and rejected in science in the light of evidence: these rules need to be designed to expose theories to the maximum risk of empirical refutation. Conventionalist stratagems, in particular, are to be banned. Faced by a refutation, a theory may be modified so as to overcome the refutation, but only if the modification increases the empirical content, the degree of falsifiability, of the theory – the modified theory predicting more, excluding more potentially falsifying observational statements, than before. Scientists should always strive to put forward theories that say as much as possible about the empirical world, that expose themselves to the greatest risk of refutation, that have the highest possible degree of falsifiability. The supreme methodological principle of science, for Popper, "says that the other rules of scientific procedure must be designed in such a way that they do not protect any statement in science against falsification" (*L.Sc.D.*, p. 54).

Even though theories cannot be verified, they can be "corroborated". For Popper, corroboration is a measure of how well a theory has stood up to attempts to refute it. If a highly falsifiable theory has survived an onslaught of severe testing, then it has proved its worth. It deserves to be taken more seriously than an untested theory, or an unfalsifiable speculation.

According to Popper, then, science makes progress by means of wild imagining, bold guesswork, on the one hand, controlled by ferocious attempted empirical refutation on the other hand.

L.Sc.D. was influenced by the thought of the Vienna Circle, but also differs from, and is highly critical of, some of the main tenets of the Circle. Logical positivism sought to demarcate the meaningful from the meaningless, only those propositions capable of being verified being meaningful, the hope being that all meaningful factual propositions would be scientific. Popper stressed that any such criterion would condemn scientific theories to being meaningless, since they cannot be verified. Popper's demarcation problem differed from that of the positivists. For Popper, as we have seen, the problem was to demarcate science from non-science (pseudo-science and metaphysics); falsifiability,

not verifiability, is the key requirement; but non-scientific, metaphysical theories, though neither verifiable nor falsifiable, may nevertheless be entirely meaningful, and may even have a fruitful role to play in the development of science. Metaphysical theses, such as atomism, may suggest, and may (as a result of acquiring precision) be transformed into, falsifiable scientific theories.

Does *L.Sc.D.* succeed in solving its two basic problems? Three great merits of the book are its originality, its clarity, and its tight structure: everything devolves from the key idea of falsifiability. This makes the book especially open to criticism, and to improvement. In the end, the book fails to solve its basic problems, due to its treatment of simplicity.

Popper claims that the more falsifiable a theory is, so the greater its degree of simplicity. (There is a second method for assessing degrees of simplicity, in terms of number of observation statements required to falsify the theories in question, but Popper stresses that if the two methods clash, it is the first that takes precedence.) It is easy to see that Popper's proposal fails. Given a reasonably simple scientific theory, T , one can readily increase the falsifiability of T by adding on an independently testable hypotheses, h_1 , to form the new theory, $T + h_1$. This new theory will be more falsifiable than T but, in general, will be drastically less simple. And one can make the situation even worse, by adding on as many independently testable hypotheses as one pleases, h_2, h_3 and so on, to form new theories $T + h_1 + h_2 + h_3 + \dots$, as highly empirically falsifiable and as drastically lacking in simplicity, as one pleases. Thus simplicity cannot be equated with falsifiability.

And there is a further, even more devastating point. Popper's methodological rules favour $T + h_1 + h_2 + h_3$ over T , especially if h_1, h_2 and h_3 have been severely tested, and corroborated. But in scientific practice, $T + h_1 + h_2 + h_3$ would never even be considered, however highly corroborated it might be if considered, because of its extreme lack of simplicity or unity, its grossly *ad hoc* character. There is here a fundamental flaw in the central doctrine of *L.Sc.D.*

Later, in *Conjectures and Refutations* (1963), Popper put forward a new methodological principle which, when added to those of *L.Sc.D.*, succeeds in excluding theories such as $T + h_1 + h_2 + h_3$ from scientific consideration. According to Popper, a new theory, in order to be acceptable, "should proceed from some *simple, new, and powerful, unifying idea* about some connection or relation (such as gravitational attraction) between hitherto unconnected things (such as planets and apples) or facts (such as inertial and gravitational mass) or new "theoretical entities" (such as field and particles)" (p. 241). $T + h_1 + h_2 + h_3$ does not "proceed from some *simple, new and powerful, unifying idea*" and is to be rejected on that account, even if more highly corroborated than T .

But the adoption of this "requirement of simplicity" (as Popper calls it) as a basic methodological principle of science has the effect of permanently excluding from science all *ad hoc* theories (such as $T + h_1 + h_2 + h_3$) that fail to satisfy the principle, however empirically successful such theories might be if considered. This amounts to assuming permanently that the universe is such that no *ad hoc* theory, that fails to satisfy Popper's principle of simplicity, is true. It amounts to accepting, as a permanent item of scientific knowledge, the substantial metaphysical thesis that the universe is non-*ad hoc*, in the sense that no theory that fails to satisfy Popper's principle of simplicity is true. But this clashes with Popper's criterion of demarcation: that no unfalsifiable, metaphysical thesis is to be accepted as a part of scientific knowledge.

It is, in fact, important that Popper's criterion of demarcation is rejected, and the metaphysical thesis of non-*ad hocness* is explicitly acknowledged to be a part of scientific knowledge. The thesis, in the form in which it is implicitly adopted at any given stage in the development of science, may well be false. Scientific progress may require that it be modified. The thesis needs to be made explicit, in other words, for good Popperian reasons, namely, so that it can be critically assessed, and perhaps improved. As long as Popper's demarcation criterion is upheld, the metaphysical thesis must remain implicit, and hence immune to criticism.

Popper's falsificationism can be modified, however, so that substantial metaphysical theses, implicit in methods that exclude

ad hoc theories, are made explicit within science, and are thus rendered available to critical scrutiny and revision (as we shall see in chapters 4, 5, 10 and 12): see also my (1974 and 1998).

On publication, *L.Sc.D.* achieved a certain impact; it was quite widely reviewed and discussed, and led to Popper being invited to give lectures, in England, Denmark and elsewhere. Popper himself later claimed that his criticisms of logical positivism led eventually to the downfall of that doctrine.

Having dealt with the methodology of the natural sciences, Popper turned his attention again to what had long been of concern to him, the intellectual defects of Marxism, and the philosophy of the social sciences. But before he could get very far with that work, Popper was offered a lectureship at the University of Canterbury, New Zealand. He accepted, and Popper and his wife left Vienna for New Zealand early in 1937.

For some years Popper had been privately highly critical of policies of socialists in Germany and Austria for playing into the hands of the Fascists and Hitler. This was due, in Popper's view, to the harmful influence of Marxism. But he had kept these criticisms to himself, as he felt any public criticism could only weaken the forces opposing Hitler. Then, in March 1938, Hitler occupied Austria, and Popper felt all grounds for restraint had disappeared. He decided to put his criticisms of Marxism, and his views on the social sciences, into a publishable form.

He began work on what was to become *The Poverty of Historicism*. But then, unexpectedly, sections on essentialism, and on totalitarian tendencies in Plato, grew and grew (driven by the desperation of the times), and Popper found he had a new work on his hands: it became what is perhaps his best known, most influential and greatest work, *The Open Society and Its Enemies* (1945). Without referring anywhere to Hitler or Stalin, the book is, nevertheless, an urgent and passionate investigation into the problem and threat of totalitarianism, whether of the right or left. It seeks to understand what the appeal of totalitarianism can be, and why it should have come to be such a threat to civilization. Popper regarded the writing of the book as his contribution to the war effort.

In *The Open Society and Its Enemies (O.S.E.)*, Popper argues that a fundamental problem confronting humanity is that of moving from a closed, tribal way of life to an open society. The closed society is a society that has just one view of the world, one set of values, one basic way of life. It is a world dominated by dogma, fixed taboos and magic, devoid of doubt and uncertainty. The open society, by contrast, tolerates diversity of views, values and ways of life. In the open society learning through criticism is possible just because diverse views and values are tolerated. For Popper, the open society is the civilized society, in which individual freedom and responsibility, justice, democracy, humane values, reason and science can flourish.

But moving from the closed to the open society imposes a great psychological burden on the individuals involved, "the strain of civilization". Instead of the security of the tribe, organic, dogmatic and devoid of doubt, there is all the uncertainty and insecurity of the open society, the painful necessity of taking personal responsibility for one's life in a state of ignorance, the lack of intimacy associated with the "abstract society" in which individuals constantly rub shoulders with strangers. This transition, from the closed to the open society is, for Popper, "one of the deepest revolutions through which mankind has passed" (*O.S.E.*, vol. 1, p. 175). Many cannot bear the burden of freedom and doubt, and long for the false security and certainties of the closed society. In particular, some of the greatest thinkers of western civilization have given into this temptation and have, in one way or another, urged upon long suffering humanity a return to something like a closed society under the guise of Utopia. This is true of Plato and Aristotle; and it is true, in more recent times, of Hegel and Marx. The lure of totalitarianism is built deep in our history and traditions.

The revolutionary transition from closed to open society first occurred, according to Popper, with the "Great Generation" of ancient Athens in the fifth century BC. Those to be associated with the birth and affirmation of the open society include Pericles, Herodotus, Protagoras, Democritus, Alcidas, Lycophon, Antisthenes and, above all, Socrates.

It is from Plato, especially, that we learn of Socrates' passionate scepticism, his searching criticism of current beliefs and ideals, his conviction that first one must acknowledge one's own ignorance before one could hope to acquire knowledge and wisdom. But Socrates, Popper argues, was ultimately betrayed by Plato. The greatest advocate of the open society became, in Plato's *Republic*, the spokesman for a return to a closed society.

Popper's devastating account of Plato's "propaganda" for the closed society, in bald outline, amounts to this. Deeply disturbed by the democracy, and the beginnings of the open society, in contemporary Athens, Plato came to fear all social change as embodying decay and corruption. Synthesizing elements taken from Parmenides, from the Pythagoreans, and from Socrates, Plato turned these fears into an entire cosmology and social theory. Every kind of material object has its perfect copy, its ideal representation, as a Form in a kind of Platonic Heaven (Plato's famous theory of Forms). These Forms initiated the material universe by printing themselves on space, thus producing initial material copies. But, as time passes, copies of copies gradually become more and more corrupt, further and further removed from their ideal progenitors. And this is just as true in the social and moral sphere as the material. The primary task for the rulers of society is to arrest all social change, and try to keep society resembling, as far as possible, the ideal Forms of order, justice and the Good. Most people know only of imperfect material things; but a very few philosophers, as a result of studying mathematics (which enables us to acquire knowledge of abstract, perfect objects and not just their imperfect material copies), are able to come to see, intellectually, the Forms, and eventually the supreme Form of the Good (represented as the sun in Plato's famous myth of the cave in *The Republic*). Enlightened philosophers alone have seen the Form of the Good; they alone know what ideal form society should take, and how it can be protected from the corrupting effects of change. Philosophers, then, must rule, aided by guardians, a class of soldiers or police, who ensure that the rest of the population obeys the strictures of the ruling philosophers. Plato's republic is a nightmarish totalitarian, closed society, rigidly ordered, individual liberty, freedom of expression and discussion,

art, democracy and justice ruthlessly suppressed. But Plato presents all this with great subtlety, with a kind of twisted logic, so that ostensibly he is arguing for a just, wise and harmonious society, one of legal and moral perfection. Popper even suggests that Plato wrote *The Republic* as a kind of manifesto, to aid his adoption as philosopher-ruler.

Popper's two big enemies of the open society are Plato (volume 1 of *O.S.E.*), and Marx (volume 2). Both uphold versions of historicism – the doctrine that history unfolds according to some fixed pattern, to some rigid set of laws of historical evolution. Plato, as we have seen, was a pessimistic historicist: historical change involves decay and degeneration, and all that enlightened philosopher-rulers can do is arrest change somewhat. Marx, by contrast, is an optimistic historicist: historical development will eventually result in socialism and freedom.

Popper traces a direct link from Plato to Marx, via Aristotle and Hegel. Prompted in part by his biological interests, Aristotle modified Plato's doctrine of the Forms so that it could give an account of biological growth and development. Aristotle inserts a Platonic Form into each individual object so that it becomes the essence of that object, an inherent potentiality which the object, through movement, change or growth, strives to realize. Thus the oak tree is inherent as a potentiality in the acorn. Germination and growth are to be understood as the acorn striving to realize its potentiality, thus becoming an oak tree.

In short, Aristotle modifies Plato's doctrine of the Forms so that the Form ceases to be the perfect copy of an object from which the object can only decay, and becomes instead an inherent potentiality which the object strives to realize. This modification potentially transforms Plato's pessimistic historicism of inevitable decay into an optimistic historicism of social growth, development and progress. But not until Hegel did anyone fully exploit Aristotelianism in this way.

Popper depicts Hegel as a complete intellectual fraud. He agrees with Schopenhauer's verdict: "Hegel, installed from above, by the powers that be, as the certified Great Philosopher, was a flat-headed, insipid, nauseating, illiterate charlatan, who reached the pinnacle of audacity in scribbling together and dishing up the

craziest mystifying nonsense" (quoted in *O.S.E.*, vol. 2, pp. 32-3). Hegel's great idea was to depict history as the process of Spirit, the Aristotelian essence and potentiality of the State and the Nation, striving to realize itself through war and world domination. Taking over and corrupting the antinomies of Kant's *Critique of Pure Reason*, Hegel depicted history as a kind of pseudo rational or logical dialectical process, *thesis* giving way to *antithesis*, which then results in *synthesis*. What matters is not individual liberty or democracy, but rather the triumph of the strongest State on the stage of history, its inner essence interpreted and directed by the Great Leader by means of dictatorial power.

Despite (or because of) his intellectual fraudulence, Hegel exercised – Popper argues – a powerful influence over the development of subsequent nationalist, historicist and totalitarian thought, of both the extreme right and the extreme left. Both Hitler and Stalin stumble onto the world stage out of Hegel, Popper implies (although neither is mentioned by name in *O.S.E.*, as indicted above). In particular, Hegel exercised a powerful and corrupting influence on Karl Marx.

For Popper, Marx is in a quite different category from Hegel. Popper pays tribute to Marx's sincerity, his humanitarianism, his intellectual honesty, his hatred of moralizing verbiage and hypocrisy, his sense of facts and his sincere quest for the truth, his important contributions to historical studies and social science, his burning desire to help the oppressed. Nevertheless Marx is one of the most dangerous enemies of the open society, his thought disastrously corrupted by its Hegelian inheritance.

In a well known passage in *Capital*, Marx declared that Hegel "stands dialectics on its head; one must turn it the right way up again" (quoted in *O.S.E.*, vol. 2, p. 102). And in another passage, Marx declared "It is not the consciousness of man that determines his existence – rather, it is his social existence that determines his consciousness" (quoted in *O.S.E.*, vol. 2, p. 89). Whereas, for Hegel, an idealist, history is the dialectical development of ideas, for Marx history is determined by the dialectical development of material processes, in particular those associated with the means of production. Distinct historical phases – pre-feudal, feudal, capitalistic, post-revolutionary socialist – owe their existence to

distinct phases in the means of production, and the social arrangements these phases generate. Each phase leads, as a result of inevitable dialectical processes, to its own destruction and the creation of the next phase. Thus capitalism concentrates wealth and ownership of the means of the production into fewer and fewer hands until, eventually, the workers unite, overthrow the capitalists and establish socialism. The historical processes of dialectical materialism work themselves out through class struggle, classes and the conflicts between them being determined by the means of production. It is the laws determining the evolution of the economic base that decide the path of history; ideas, democratic and legal institutions form an ideological superstructure, which reflects the economic base and the interests of the dominant class, but is powerless to influence the path of history. Marx condemned as "Utopian" those socialists who sought to bring about the revolution by means of political policies and plans. The proper "scientific" approach to bringing about socialism is, first, to discover the dialectical laws governing the evolution of the economic base of society, and then to help this evolution along, in so far as this is possible, thus speeding up the coming of the final, inevitable socialist revolution.

Popper argues that a number of elements of Marxist thought are of value, if not taken too far. There is the idea that the social cannot be reduced to the psychological, sociology not being reducible to psychology. There is the thesis that much of history has been influenced by class struggle, and the idea that the means of production, economic circumstances, play an important role in influencing the development of other aspects of social and cultural life, even something as apparently remote from economic conditions as mathematics. Above all, there is the recognition and depiction of the appalling conditions of life of the poor in the unrestrained capitalist conditions of Marx's time, and the recognition, too, of the hypocrisy of much of the morality, the legal system and the politics of those times. Having described Marx's account of the working conditions of children as young as 6 years, Popper writes: "Such were the conditions of the working class even in 1863, when Marx was writing *Capital*; his burning protest against these crimes, which were then tolerated, and sometimes

even defended, not only by professional economists but also by churchmen, will secure him forever a place among the liberators of mankind" (*O.S.E.*, vol. 2, p. 122).

But these good points are, for Popper, more than counterbalanced by the dreadful defects, most of which stem from Marx's historicism, inherited from Hegel. For the central tenet of Marxism is the idea that the laws of dialectical materialism determine the evolution of the means of production, and this in turn determines the evolution of everything else, from class struggle to culture, religion, the law and politics. But this is manifestly false. For one thing, there is a two-way interaction between economic conditions and ideas; eliminate scientific and technological ideas, and the economy would collapse. For another, ideas can themselves influence the course of history, Marxism itself being an example. Historical predictions made by Marx, on the basis of his economic historicism, have been falsified by subsequent historical events. The Russian revolution is, for example, entirely at odds with Marx's theory, as is the way in which the unrestricted capitalism of Marx's time has subsequently become both more economically successful and more just and humane as a result of diverse political interventions. Marx's economic historicism is not just false; it is pseudo-scientific. Only for exceptionally simple systems, such as the solar system, is long-term prediction, based on scientific theory, possible. In the case of social systems, incredibly complex and open to the influence of a multitude of unpredictable factors, the idea that science should be able to deliver long-term predictions is hopelessly unwarranted. Marx's historicism leads him to turn good points into bad ones by exaggeration. "The history of all hitherto existing society is a history of class struggle" (quoted in *O.S.E.*, vol. 2, p. 111) is a good point if "all" is not taken too seriously, but as it stands is an oversimplification and exaggeration; it ignores, for example, power struggles within the ruling class. Again, Marx was surely right to see legal and political institutions of his time as being biased in the direction of the interests of the ruling classes; but he was wrong to condemn all legal and political institutions as inevitably having this function, as his economic historicism compelled him to do.

For Popper, the most damaging feature of Marx's historicism has to do, perhaps, with the severe limitations that it places on the power of politics, on the capacity of people to solve social problems. Marx is famous for his eleventh thesis on Feuerbach: "The philosophers have only interpreted the world in various ways; the point however is to change it" (quoted in *O.S.E.*, vol. 2, p. 84). But Marx's economic historicism leads immediately to a severely restricted view as to what political intervention can achieve. In *Capital* he declares: "When a society has discovered the natural law that determines its own movement, . . . even then, it can neither overleap the natural phases of its evolution, nor shuffle them out of the world by a stroke of the pen. But this much it can do; it can shorten and lessen its birth-pangs" (quoted in *O.S.E.*, vol. 2, p. 86). Just those actions which were to improve the unrestrained capitalism of Marx's time beyond all recognition, namely political intervention and the actions of trade unions, are discounted at the outset by Marx's economic determinism as necessarily impotent. Political planning and policy making for socialism is condemned by Marx, in line with his central doctrine, as inherently inefficacious and Utopian. One disastrous consequence of this was that when Marxists gained power in Russia, they found their literature contained no guidelines as to how to proceed. Another disastrous consequence was that Marxism, blind to the potency of political power, failed to anticipate the dangers inherent in handing over power to political leaders after the revolution, dangers which, after the Russian revolution, became all too manifest.

The full force of Popper's criticism is devoted, however, to the central argument of *Capital* – an argument which seeks to establish the inevitable downfall of capitalism and the triumph of socialism. Popper presents Marx's arguments as having three steps, only the first of which is elaborated in *Capital*. The first step argues that an inevitable increase in the productivity of work leads to the accumulation of more and more wealth in the ruling class, and the greater and greater poverty and misery of the working class. The second step then argues that all classes will disappear except for a small, wealthy ruling class and a large impoverished working class, this situation inevitably leading to a revolution. The third step argues that the revolution will result in the victory of the

working class, which in turn will result in the withering away of the state and the creation of socialism.

Popper demonstrates that none of these steps is inevitable by showing that alternative developments are entirely possible and, in many cases, have actually happened after Marx wrote *Capital*. Even if there is a tendency under capitalism for the means of production and wealth to be concentrated in fewer and fewer hands (as the first step assumes), the state can intervene to counteract this tendency by such means as taxation and death duties. And as far as the increasing poverty of workers is concerned, this can be counteracted by the formation of trade unions, by collective bargaining backed up by strikes. The brutal, unrestricted capitalism of Marx's time has been transformed since out of all recognition by just such interventionist methods. And Popper makes analogous, decisive points to demolish the second and third steps of Marx's argument. Even if the ruling class did become increasingly wealthy and the working class increasingly poor (as the second step assumes), this does not mean that all classes but these two would necessarily disappear, since landowners, rural workers, and a new middle class may well exist, given Marx's assumptions. And even if violence breaks out, this does not mean it would necessarily constitute the social revolution, as envisaged by Marx. And finally, even if it is granted that the workers unite and overthrow the ruling class (as the third step assumes), this does not mean that a classless society and socialism would necessarily result. It is all too easy to suppose that the new political leaders would seize and hold onto power, justifying this by exploiting and twisting the revolutionary ideology, and by invoking the threat of counter-revolutionary forces. And many other possible outcomes can be envisaged. It is in fact implausible to suppose that the victory of the working class would mean the creation of a classless society, and hence the withering away of the state. (This bald summary does not begin to do justice to the cumulative force of Popper's argument.)

Marx, as we have seen, condemned planning for socialism as Utopian; and in a sense Popper agrees. Popper distinguishes two kinds of social planning or intervention, which he calls Utopian and piecemeal social engineering. Utopian social engineering

seeks to attain an ideal social order, such as socialism, by bringing about holistic changes in society; such an approach is, Popper argues, doomed to failure. Piecemeal social engineering, by contrast, searches for and fights against "the greatest and most urgent evils of society": this is the approach that Popper advocates (*O.S.E.*, vol 1, ch. 9). Subsequently, during the course of criticizing Marx, Popper points out that piecemeal social engineering can take the form either of state intervention, or of the creation of legal, institutional checks on freedom of action. The latter is to be preferred, Popper argues, as the former carries with it the danger of increasing the power of the state (*O.S.E.*, vol. 2, pp. 129-133).

There is very much more to Popper's *O.S.E.* than the above indicates. Central to the book is the idea that reason is a vital component of the open society, reason being understood as "critical rationalism", arrived at by generalizing Popper's falsificationist conception of scientific method. For Popper, both scientific method, and rationality, need to be understood in social terms. Popper criticizes Karl Mannheim's sociology of knowledge for overlooking the "social aspect of scientific method" (*O.S.E.*, vol. 2, ch. 23). Popper criticizes moral historicism, oracular philosophy and the revolt against reason, and the idea that history might have a meaning (*O.S.E.*, vol. 2, chs. 22, 24 and 25 respectively). Both volumes have extensive footnotes containing fascinating discussion of a great variety of issues tangentially related to the main argument, such as the development of ancient Greek mathematics, the problem of putting an end to war, or the proper aims of a liberal education.

Popper's fiercely polemical book has provoked much controversy. His critical onslaughts against Plato, Aristotle, Hegel and Marx have been angrily repudiated by many scholars in these fields; or, much worse, just blandly ignored.

The general doctrine of historicism is expounded and criticized by Popper in his *The Poverty of Historicism (P.H.)*, first published in three parts in *Economica* in 1944 and 1945 (somewhat before *O.S.E.*) and only published in book form in 1957. Popper divides historicist views and arguments into two classes: those that hold that the methods of the social and natural scientists are quite

different (the "anti-naturalist doctrines") and those that hold they are the same or similar (the "pro-naturalist doctrines").

Anti-naturalist doctrines can be summarised like this. Generalizations, experiments, predictions and understanding have roles in social science that are radically different from those they have in physics. Social phenomena exhibit novelty, complexity, and a holistic aspect that is lacking in physical phenomena. These differences ensure that historicist social science, predicting in more or less rough outline the evolution of society, must employ methods that differ from those of natural science.

Pro-naturalist doctrines of historicism make much of the success of long-term predictions in astronomy. Just as states of the solar system can be predicted by natural science far into the future, so too historicist social science ought to be able to predict states of society far into the future. Such predictions will, however, employ social laws of succession, laws which specify how one characteristic phase of social development give way to a subsequent phase.

Popper effectively criticizes the anti-naturalist doctrines of historicism. It is, however, Popper's criticism of the pro-naturalist standpoint that is the really important nub of the book. Historicist laws of successions are not laws at all, as these are understood in physics. They are *trends*. And "*trends are not laws*" (P.H., p. 115). A law provides a causal explanation of an event when the law *plus initial conditions* imply that the event occurs. Whenever a succession of causally connected events occur in our environment, such as the wind shaking a tree and causing an apple to fall to the ground, laws (usually a number of quite different laws) *plus the specification of a sequence of initial conditions* are required to predict the sequence of events. Trends can, then, be explained by means of laws, but it is always laws *plus relevant initial conditions* which provide such explanations. And the crucial point is that, given some trend, in particular a social trend, initial conditions that must continue to exist if the trend is to continue, are likely to be very many indeed, most of which will be easy to overlook. This ensures that trends, such as the growth of a population, which have persisted for centuries, may quite suddenly cease if some condition, necessary for the persistence of the trend,

ceases to exist. "The poverty of historicism", Popper declares "is a poverty of imagination" (p. 130) – the poverty of being unable to imagine that conditions, necessary for the persistence of some trend, might suddenly themselves change. And this is highly relevant to the whole idea of piecemeal social engineering, for the piecemeal engineer may seek to change just such conditions, required for the persistence of some undesirable trend.

In 1945 Popper was appointed to a readership in Logic and Scientific Method at the London School of Economics; he took up the appointment in 1946, and was promoted to a personal professorship in 1949. Initially the only philosopher at the L.S.E., Popper was subsequently joined by J. O. Wisdom in 1948, Joseph Agassi in 1957 (who left in 1960), John Watkins in 1958, W. W. Bartley III and Imre Lakatos in 1960, and Alan Musgrave in 1964. The Department at the L.S.E. was famous for Popper's weekly seminar. Notoriously, visiting speakers rarely succeeded in concluding the announcement of the title of their talk before being interrupted by Popper. Each speaker was subjected to a devastating critical attack by Popper, almost sentence by sentence; quite often, the subject of the seminar would be continued a week later. The seminars were always dramatic, sometimes farcical, but nevertheless created an overwhelming impression of Popper's passionate determination to get at the truth, even if conventions of politeness and good manners had to be sacrificed.

In 1963 Popper published *Conjectures and Refutations (C.R.)*, a collection of essays restating, extending and applying his views on scientific method, philosophy and rationality. This is perhaps the best introduction to Popper's work. Here is a quick survey of some of the items in the book.

In the Introduction Popper makes a number of important epistemological points. He notes the widespread tendency to believe in the false doctrine that truth is manifest – readily available and easy to come by. When truth turns out not to be so easy to obtain, epistemological optimists become pessimists, and deny that knowledge is possible at all, or resort to conspiracy theories to account for the inaccessibility of the truth. Both Descartes and Bacon are famous for their anti-authoritarian stance in epistemological matters; and yet, Popper points out, there is an

unnoticed implicit authoritarianism in their views. For Bacon, and for the empiricists who followed him, the senses are authoritative sources of knowledge; for Descartes, and for the rationalists who followed him, reason is the authoritative source of knowledge. Popper, of course, argues against the idea that conjectural knowledge has any authoritative source.

In chapter two, Popper argues that philosophical problems have their roots in science and mathematics; and he argues against the Wittgensteinian view that philosophical problems are pseudo-problems that arise when ordinary language is misused.

In chapter three Popper distinguishes three views concerning human knowledge: essentialism, which holds that science can grasp the ultimate essence of things; instrumentalism, which holds that scientific theories are merely instruments for the prediction of observable phenomena; and realism, which holds that science puts forward falsifiable conjectures about aspects of reality that often go beyond what is observable. Popper criticizes the first two views, and defends the third view.

In chapter five Popper gives a magnificent account of the Presocratic philosophers – Thales, Anaximander, Anaximenes, Heraclitus, Parmenides – as proposing and critically assessing successive theories about the origins and ultimate constituents of the universe, and about the problem of how to understand change. The Presocratics, Popper argues, almost unintentionally created critical rationality, the tradition of proposing bold conjectures which are then subjected to criticism – a tradition that led eventually to modern science.

In chapter eight Popper tackles the problem of how philosophical or metaphysical doctrines can be rationally assessed given that they cannot be empirically falsified, like scientific theories. His solution is that philosophical doctrines can be assessed from the standpoint of the problems that they are intended to solve; even though irrefutable, they can nevertheless be criticized from the standpoint of the problems they seek to solve.

In chapter ten Popper restates and, as we have seen above, develops somewhat his falsificationist conception of scientific method. In this chapter Popper formulates and tries to solve what has subsequently come to be known as *the problem of*

verisimilitude: What can we mean by scientific progress if science advances from one false theory to another? Popper's solution is that, given two theories, T_1 and T_2 , even though both are false, nevertheless T_2 say, may be closer to the truth than T_1 . Suppose, for example, that T_2 implies everything true that T_1 implies and more besides, but T_2 does not imply anything false that T_1 does not imply. Granted this, there is a perfectly good sense in which T_2 can be said to be "closer to the truth" than T_1 , and thus an advance over T_1 . Unfortunately it was subsequently shown by Tichy (1974) and Miller (1974) that this proposed solution to the problem does not work. If T_2 has more true implications than T_1 does, then T_2 , necessarily, has some false implications which T_1 does not have. Popper's requirements for T_2 to be closer to the truth than T_1 , when both are false, cannot be satisfied.

Chapter fifteen provides an exposition and decisive criticism of dialectic reasoning: it is thus an adjunct to the criticisms of Hegel and Marx to be found in *O.S.E.*.

One of the themes running through *C.R.*, and through much of Popper's subsequent writings, is that the proper task of philosophy is to tackle, in an imaginative and critical way, real, fundamental problems having their roots outside philosophy in science, politics, art, life. This Popperian conception of philosophy stands in sharp contrast both to the pomposities and obscurities of much so-called "continental" philosophy, and to the poverty and aridity of philosophy in the so-called "analytic" tradition, restricted to ordinary language analysis, the analysis of meaning. Popper has fought against both rival conceptions of philosophy, and has sought to put into practice his own critical rationalist, problem-solving conception of philosophy. His first four books are exemplary in this respect, and have undoubtedly exercised an enormous, healthy influence on much subsequent philosophy, even though this influence has often not been acknowledged. A basic impulse behind these works might almost be summed up in a stray remark tossed out in *O.S.E.*, vol. 2 (p. 59): "We have to learn the lesson that intellectual honesty is fundamental for everything we cherish".

It is important to appreciate the existence of a central backbone of argument running through these four books. In *L.Sc.D.*, as we

have seen, Popper argues that all scientific knowledge is irredeemably conjectural in character, it being impossible to verify theories empirically. Science makes progress by proposing bold conjectures in response to problems, which are then subjected to sustained attempted empirical refutation. This falsificationist conception of scientific method is then generalized to form Popper's conception of (critical) rationality, a general methodology for solving problems or making progress. As Popper puts it in *L.Sc.D.* "inter-subjective *testing* is merely a very important aspect of the more general idea of inter-subjective *criticism*, or in other words, of the idea of mutual rational control by critical discussion" (p. 44, note 1*). But in order to make sense of the idea of *severe testing* in science, we need to see the experimentalist as having at least the germ of an idea for a rival theory up his sleeve (otherwise testing might degenerate into performing essentially the same experiment again and again). This means experiments are always *crucial experiments*, attempts at trying to decide between two competing theories. Theoretical pluralism is necessary for science to be genuinely empirical. And, more generally, in order to *criticize* an idea, one needs to have a rival idea in mind. Rationality, as construed by Popper, requires plurality of ideas, values, ways of life. Thus, for Popper, the rational society *is* the open society. Given pre-Popperian conceptions of reason, with their emphasis on proof rather than criticism (and associated plurality of ideas), the idea that the rational society is the open society is almost a contradiction in terms. There is thus a very close link between *L.Sc.D.*, on the one hand, and *O.S.E.*, *P.H.* and *C.R.* on the other. And the direction of argument does not go in just one direction, from *L.Sc.D.* to *O.S.E.*: it goes in the other direction as well. For in *O.S.E.* (vol. 1, ch. 10), Popper argues that rationality, and scientific rationality as well, need to be conceived of in social and institutional terms (and the argument is echoed in *P.H.*, in connection with a discussion about the conditions required for scientific progress to be possible). *O.S.E.*, *P.H.* and *C.R.* illuminate and enrich the doctrines of *L.Sc.D.*.

Above, in connection with the discussion of *L.Sc.D.*, it was argued that Popper's falsificationism ultimately fails, because of its failure to exclude highly falsifiable but grossly *ad hoc* theories

from science. The scientific enterprise is obliged to conjecture that the universe is more or less comprehensible, having some kind of unified dynamic structure, only those theories being tentatively accepted which satisfy (a) empirical considerations, and (b) considerations having to do with simplicity, unity, comprehensibility. As science proceeds, we improve our (conjectural) knowledge of the kind of comprehensible unity which may exist in nature; the *aim* of science improves, and with it the *methods* of science. There is, in other words, a kind of positive feedback between improving knowledge and improving aims and methods, improving knowledge about how to improve knowledge. Science adapts its nature to what it finds out about the nature of the universe (which helps to account for the almost explosive growth of scientific knowledge).

This "evolving-aims-and-methods" view of science modifies quite considerably Popper's falsificationism. When generalized, it leads to an "evolving-aims-and-methods" view of rationality which in turn modifies quite considerably Popper's critical rationalism. These modifications, if adopted, have far reaching implications for central doctrines of Popper's *L.Sc.D.*, *O.S.E.*, *P.H.* and *C.R.*: see chapters 4, 5 and 10, and my (1984, 1998, 2001 and 2010).

Work published by Popper after *C.R.*, though containing much of great value, is not, perhaps, in quite the same league as that of his first four books. Much of this work restates, extends and further applies earlier ideas. Where Popper's subsequent work launches forth in new directions, these are not always well chosen. Battles against subjectivity, anti-realism and physical determinism lure Popper into defending opposing views that are exaggerated, sometimes, almost to the point of absurdity. A subtle shift of perspective, of allegiance, can be discerned as we move from Popper's earlier to his later work. In his early work, Popper speaks up on behalf of humanity, on behalf of any concerned person of good will, and against those traditional "great thinkers" and "experts" who threaten to beguile us and lead us to disaster. In his later work, the allegiances have shifted; now Popper speaks up on behalf of great science and great scientists, and against fraudulent academics, mostly philosophers and social scientists.

In 1970 there appeared *Criticism and the Growth of Knowledge*, edited by Imre Lakatos and Alan Musgrave, the fourth volume of the proceedings of a conference on philosophy of science held in London in 1968. This volume is devoted to a comparison of the views of Thomas Kuhn and Popper on the philosophy of science, and contains contributions from Kuhn, Popper, Watkins, Toulmin, Lakatos, Feyerabend, and others. In his contribution Popper praises Kuhn for having discovered normal science, science which takes some "paradigm" for granted and devotes itself to puzzle solving. Popper points out that he had himself made the same discovery over thirty years earlier, as recorded in the preface to *L.Sc.D.*. But the normal scientist "has been badly taught. He has been taught in a dogmatic spirit: he is the victim of indoctrination". Normal science is "a danger to science and, indeed, to our civilization" (p. 53).

In 1972 Popper published a second collection of his essays entitled *Objective Knowledge*. One of the essays makes the good point that common sense tends to combine two incompatible theses, common sense realism, and the epistemological view that knowledge comes flooding into our minds via the senses, rather like water being poured into a bucket – a view which Popper's dubs "the bucket theory of the mind". Popper argues that these two theses clash, and that philosophers, registering this clash, have all too often held onto the bucket theory and rejected realism. But this, Popper argues, is exactly the wrong thing to do; one should hold onto realism, and reject the bucket theory.

Much of the rest of the book is devoted to developing and defending Popper's three-world view. There are, according to this view, three worlds: the physical world (world 1), the psychological or mental world (world 2), and the world of objective theories, propositions, arguments and problems (world 3). World 3 interacts with world 1 via world 2. Popper argues that this interaction is demonstrated by the fact that scientific theories lead to new technology, world 1 phenomena, which would not exist were it not for the prior development of world 3 theories. Popper puts world 3 into a biological and evolutionary context: like the webs, nests and dams created by spiders, birds and beavers, so too world 3 is our

creation but, once created, it acquires an objective existence independent of us.

This theme is continued in Popper's contribution to *The Self and Its Brain (S.B.)* a book written with the neurologist, John Eccles. In this work, Popper develops a sustained argument in support of interactionism and his three-worlds view, and criticizes materialism, physicalism, and the thesis that the physical world is (causally) closed. There is also an interesting chapter on the history of the mind-body problem, in which Popper argues for the questionable thesis that the problem was recognized independently of, and before the arrival of, anything like the modern scientific view of the world.

What is one to make of this three-worlds view? Popper is surely right to hold that the contents of theories need to be distinguished from their linguistic forms (and from the causal effects these linguistic forms can have on appropriately educated brains). Popper is also right, surely, to stress that, in order to make human sense of human action we need to attend to the contents of theories. But it is quite another matter to argue, as Popper does, that world 3 entities, such as contents of theories, exist as full-blooded, almost Platonic entities, poltergeistic intellectual objects capable of influencing material phenomena via their influence on conscious minds. Popper overlooks or ignores the possibility that the material world may be *causally* closed but not *explanatorily* closed. He overlooks, that is, the possibility that physical phenomena, such as those associated with human actions and human technology, can be explained and understood in two distinct (but perhaps interdependent) ways: (1) physically and causally, in terms of physical theory, and (2) "personalistically", in terms of the intentions, plans and ideas of people. Such a view would hold that personalistic explanation is compatible with, but not reducible to, physical explanation. This view would give to the contents of theories a vital role in the (personalistic) explanation of human actions and the development of technology, without in any way undermining the existence, in principle, of a purely physical, causal explanation of physical phenomena associated with human action and technology (see Maxwell, 2001).

Popper insists that his world 3 entities differ from Plato's Forms in that they are man-made, consist of theories, including false theories, and problems, rather than reified concepts or essences, and there is no suggestion that world 3 objects can be known with certainty (*S.B.*, pp. 43-4). But even if Popper's world 3 entities do not have implausible epistemological Platonic features, they most certainly have highly implausible ontological and causal Platonic features, in that they have causal effects on the material world (via their influence on conscious minds). That the elderly Popper should espouse such an implausible Platonic doctrine almost seems like Plato's revenge for the youthful Popper's onslaught against him.

Chapter three of *Objective Knowledge* is called "Epistemology Without a Knowing Subject". Despite the title, Popper does not altogether neglect the personal dimension of the search for knowledge. What he does argue is that subjective knowledge is irrelevant to the study of scientific knowledge, only knowledge construed in objective, impersonal, world 3 terms being important. But this downplays the point that all of objective knowledge, stored in books and libraries, is of value only in so far as it is understood and used by people. Albert Einstein once remarked: "Knowledge exists in two forms – lifeless, stored in books, and alive in the consciousness of men. The second form of existence is after all the essential one; the first, indispensable as it may be, occupies only an inferior position" (Einstein, 1973, p. 80). Einstein's priorities seem saner than later Popper's. And altogether saner, more humane and down-to-earth than elderly Popper's spooky world 3 objects is the viewpoint of the more youthful Popper of *O.S.E.*, which sees science and reason in personal, social and institutional terms, without any appeal being made to ghostly, quasi-Platonic Forms.

In 1974 Popper became the fourteenth subject of *The Library of Living Philosophers*, edited by P. A. Schilpp. This two-volume work opens with Popper's "Intellectual Autobiography", subsequently published independently as *Unended Quest* (1976), continues with descriptive and critical papers on diverse aspects of Popper's work, by Quine, Putnam, Lakatos, Medawar, Watkins, Ayer, Margenau, Grünbaum, Kuhn and others, and concludes with

Popper's replies. *Unended Quest* is a fascinating book, and gives a gripping account of Popper's life-long, passionate engagement with his fundamental problems and concerns. It includes a marvellous discussion of the development of polyphonic music, and provides an account of Popper's battles with subjectivism in physics in connection with quantum theory, and with thermodynamics and the arrow of time. Popper also declares that it was he who killed logical positivism. This book, together with *C.R.*, provides the best introduction to Popper's philosophy.

In 1982-83 there appeared *Postscript to the Logic of Scientific Discovery*, a three volume work which extends and elaborates doctrines and arguments of *L.Sc.D.*, and much of which was written in the years 1951-56. The work reached the stage of proofs in 1956-57, but was abandoned because Popper suffered from detached retinas, had operations on both eyes, his sight for a time in question. It was only much later, under the editorship of W. W. Bartley III, and after some additions and rewriting, that the work finally appeared.

Volume one, entitled *Realism and the Aim of Science*, restates and elaborates Popper's earlier views and arguments concerning induction, falsification, corroboration, demarcation, realism, metaphysics and probability. At one point Popper illuminatingly contrasts how a scientific paper might be written in the style of inductivism, and in the critical, problem-solving approach of falsificationism and critical rationalism: see pages 47-51.

Volume two, entitled *The Open Universe: An Argument for Indeterminism*, sets out to refute determinism. Popper distinguishes between 'scientific' and metaphysical determinism. 'Scientific' determinism asserts that future states of physical systems can be predicted with any degree of precision by means of theories and initial conditions specified with sufficient precision: see page 36. Metaphysical determinism asserts merely that "all events in this world are fixed, or unalterable, or predetermined" (p. 8). Popper spells out an argument which, he claims, refutes scientific determinism. Even given a universe in which all events occur in accordance with a deterministic physical theory, T, nevertheless a predictor, put within an isolated system, could not predict all future states of the system with unlimited precision.

Even if such a predictor had unprecedented powers to acquire knowledge of initial conditions, and make predictions using T, nevertheless it could not acquire up-to-date information about its own state, because the attempt to do so would continually alter its state. This means it would not be possible for the predictor to predict future states of the system of which it forms a part. Popper goes on to argue against metaphysical determinism.

Although full of interesting points, there are two oddities about this discussion. First, as Popper admits, his refutation of 'scientific' determinism does not refute a second version of 'scientific' determinism which asserts that past states of physical systems can be predicted, employing prior initial conditions and physical theory. Second, Popper ignores a rather different third version of 'scientific' determinism, which asserts that the universe is such that there is a discoverable, true, physical 'theory of everything', T, which is deterministic. This version of determinism deserves to be called 'scientific' because T is asserted by it to be scientifically discoverable; furthermore, once discovered, T will be falsifiable, and hence, by Popper's own standards, scientific. It is curious that Popper, who is elsewhere (as we have seen) opposed to instrumentalism and in favour of realism, should here discuss at length a version of 'scientific' determinism which is thoroughly instrumentalistic in character, in that it makes assertions about *predictability*, and should ignore a version of 'scientific' determinism which is much more in keeping with scientific realism, in that it makes an assertion about the nature of the universe. This oversight seriously weakens Popper's argument for indeterminism.

Volume three is called *Quantum Theory and the Schism in Physics (Q.T.S.P.)*. It is concerned with quantum theory and probability, interconnected issues which preoccupied Popper, on and off, throughout his working life.

Thus in *Logic der Forschung* (1934), Popper tackled two problems concerning probability: How are probabilistic statements or theories to be interpreted? How can probabilistic theories be falsifiable given that they are in principle "impervious to falsification"? In response to the first problem, Popper defended a version of von Mises' objective, frequency interpretation of

probability. In response to the second, he insisted that probabilistic statements become falsifiable as a result of a methodological decision to treat them as falsifiable.

Logic der Forschung also devoted a chapter to problems of quantum theory. The main task is to criticize Bohr's and Heisenberg's orthodox interpretation of quantum theory (which gives equal weight to the two "complementary" pictures of particle and wave), and to provide an alternative which interprets the theory as an objective, realistic statistical theory about *particles*. Popper criticizes Heisenberg's interpretation of his uncertainty relations, which interprets these relations as placing restrictions on (simultaneous) measurement. Popper argues that these relations need to be interpreted as "scatter relations", restricting what can be predicted, and not what can be measured. Indeed, Popper argues, not only *can* we make measurements, for example, simultaneous measurements of position and momentum, that are more precise than allowed by the uncertainty relations as interpreted by Heisenberg; we *need* to do this in order to test these relations.

These issues are restated and further developed in the *Postscript*, taking into account relevant developments in quantum physics itself, such as John Bell's proof that local hidden variable versions of quantum theory cannot reproduce all the predictions of orthodox quantum theory, and experiments, such as those of Aspect, which seem to have refuted these local hidden variable theories. The main change in Popper's views is his development of his "propensity" interpretation of probability, and his application of this to quantum theory.

Popper's propensity idea is perhaps best understood in terms of an example. Consider tossing a die on a table. There is a certain probability of obtaining a six, which may or may not equal $1/6$. This is determined by such things as properties of the die (e.g. whether or not it is made of a homogeneous material), the procedure for tossing, and the properties of the table. It is this combination of properties that is, for Popper, the propensity: it is a property, determining a probability associated with some repeatable event (such as tossing the die), "of the *whole repeatable experimental arrangement*" (p. 71). In particular, then, the probabilistic statements of quantum theory can be interpreted as

attributing propensities, not to individual electrons or photons as such, but rather to electrons or photons in the context of some specific, repeatable measurement.

Popper's views on quantum theory have been criticized by Paul Feyerabend (1968-69), on the grounds that Popper fiercely criticizes Bohr but ends up defending a view very close to Bohr's. Because propensities are properties defined in terms of experimental arrangements, this means that Popper's propensity interpretation of quantum theory, just like Bohr's interpretation, brings in measurement in an essential way. Popper's reply is that propensities relate to "physical situations" which may, but need not be, experimental arrangements (Q.T.S.P., p. 71). But this reply fails in two ways. First, the probabilistic predictions of standard quantum theory *are* restricted to measurements: if these predictions are to include "physical situations" that are not measurements, then they need to be specified, and need to have specified quantum observables associated with them, so that definite probabilistic predictions may be forthcoming: Popper provides nothing of this. Second, even if Popper did extend the interpretation of quantum theory in the way just indicated, what would result would be a version of quantum theory which would reproduce most of the serious defects of the theory given Bohr's interpretation. These defects include being vague, ambiguous, *ad hoc* and non-explanatory, all resulting from the fact that the theory is made up of two incoherent parts, a quantum mechanical part, and a classical part specifying measurement or specific "physical situations".

It may, however, be possible to overcome these defects by modifying Popper's propensity version of quantum theory, so that quantum propensities determine probabilistically how quantum entities, such as electrons and photons, interact with each other (rather than with classically described, macroscopic, measuring instruments or "physical situations"). But this leads to a fully micro-realistic propensity version of quantum theory, very different from Popper's version (see chapter 13, and my 1982, 1988, 1994, 2011a). Quantum theory emerges as a theory that is about, not particles, but a new kind of probabilistic entity, the "propensiton" (as it may be called), which is neither a particle nor a wave, even though it has some features of both. Furthermore,

according to such a version of quantum theory, probabilistic transitions involve something like "wave-packet collapse" as a real physical process: for Popper, any such idea is just another part of "the great quantum muddle". But what this indicates is that here, as elsewhere in his work, Popper's ideas, even when wrong or inadequate, are nevertheless rich in fruitful suggestions and implications for further development.

After the *Postscript*, a number of collections of essays have appeared, restating and elaborating themes already indicated: *A World of Propensities* (1990), *In Search of a Better World* (1992), *The Myth of the Framework* (1994), *Knowledge and the Body-Mind Problem* (1994), *Lesson of this Century* (1997), *The World of Parmenides* (1998), and *All Life is Problem Solving* (1999).

Popper was knighted in 1965, and became a Companion of Honour in 1982. He retired from his position at the LSE in 1969. He became a Fellow of the Royal Society in 1976. Popper's wife, Hennie, died in 1985, after a long struggle with cancer. In his later years, Popper was showered with academic honours of various kinds: membership of many academic societies, honorary degrees, conferences dedicated to his philosophy, honours, medals and prizes from various sources (see Miller, 1997, pp. 403-6). Popper died a week after a serious operation, on 17th September 1994.

References

Works by Popper

Books:

Logik der Forschung: Zur Erkenntnistheorie der modernen Naturwissenschaft (Vienna: Julius Springer, 1934).

The Open Society and Its Enemies. Volume I The Spell of Plato. Volume II The High Tide of Prophecy: Hegel, Marx, and the Aftermath (2 volumes, London: George Routledge, 1945; revised edition, Princeton: Princeton University Press, 1950).

The Poverty of Historicism (London: Routledge and Kegan Paul, 1957; Boston: Beacon Press, 1957).

The Logic of Scientific Discovery (translation from the German of *Logik der Forschung* with additional footnotes and appendices, London: Hutchinson, 1959; New York: Basic Books, 1959).

Conjectures and Refutations: The Growth of Scientific Knowledge (New York: Basic Books, 1962; London: Routledge and Kegan Paul, 1963).

Objective Knowledge: An Evolutionary Approach (Oxford and New York: Clarendon Press, 1972).

The Self and Its Brain: An Argument for Interactionism, with J. C. Eccles (Berlin, New York and London: Springer-Verlag, 1977).

Die beiden Grundprobleme der Erkenntnistheorie [1930-33], edited by T. E. Henson (Tübingen: Mohr, 1979).

Unended Quest: An Intellectual Autobiography (London: Fontana/Collins, 1976).

Postscript to the Logic of Scientific Discovery, edited by W. W. Bartley, III: *Volume I: Realism and the Aim of Science* (Totowa, NJ: Rowman and Littlefield, 1982; London: Hutchinson, 1983.)

Volume II: The Open Universe: An Argument for Indeterminism (Totowa, NJ: Rowman and Littlefield, 1982; London: Hutchinson, 1982.)

Volume III: Quantum Theory and the Schism in Physics, (Boston: Unwin Hyman, 1982; London: Hutchinson, 1983).

A World of Propensities (Bristol: Thoemmes, 1990).

In Search of a Better World: Lectures and Essays from Thirty Years, translated by L. J. Bennett (London and New York: Routledge, 1992).

Knowledge and the Body-Mind Problem: In defence of interaction, edited by M. A. Notturmo (London and New York: Routledge, 1994).

The Myth of the Framework: In defence of science and rationality, edited by M. A. Notturmo (London and New York: Routledge, 1994).

The Lesson of this Century: with two talks on freedom and the democratic state, with an introduction by G. Bosetti, translated by P. Camiller (London and New York: Routledge, 1997).

The World of Parmenides: Essays on the Presocratic Enlightenment, edited by A. F. Petersen and J. Mejer (London and New York: Routledge, 1998).

All Life is Problem Solving, translated by P. Camiller (London and New York: Routledge, 1999).

Selected Periodical Publications – Uncollected

"What can Logic do for Philosophy?," *Aristotelian Society, Supplementary Volume XXII: Logical Positivism and Ethics* (1948): 141-154.

"A note on natural laws and so-called contrary-to-fact conditionals," *Mind* 58 (1949): 62-66.

"Indeterminism in Quantum Physics and in Classical Physics," *British Journal for the Philosophy of Science* 1 (1950): 117-133 and 173-195.

"The Principle of Individuation," *Aristotelian Society, Supplementary Volume XXVII: Berkeley and Modern Problems* (1953): 97-120.

Papers on "The Arrow of Time," "Irreversibility" and "Entropy," *Nature* 177 (1956): 538; 178 (1956): 382; 179 (1957): 1297; 181 (1958): 402-403; 207 (1965): 233-234; 213 (1967): 320; 214 (1967): 322.

"Probability Magic or Knowledge out of Ignorance," *Dialectica* 11 (1957): 354-372.

"The Propensity Interpretation of the Calculus of Probability, and the Quantum Theory," *Observation and Interpretation*, edited by S. Körner (London: Butterworth Scientific Publications, 1957), 65-79, 88-89.

"Irreversibility; or Entropy since 1905," *British Journal for the Philosophy of Science* 8 (1957): 151-155.

"The Propensity Interpretation of Probability," *British Journal for the Philosophy of Science* 10 (1959): 25-42.

"Philosophy and Physics," *Proceedings of the XIIth International Congress of Philosophy* (1961): 367-374.

"A Theorem on Truth-Content," *Mind, Matter and Method: Essays in Philosophy and Science in Honor of Herbert Feigl*, edited by P. Feyerabend and G. Maxwell (Minneapolis, Minnesota: University of Minnesota Press, 1966), 343-353.

"Normal Science and Its Dangers," *Criticism and the Growth of Knowledge: Proceedings of the International Colloquium in the Philosophy of Science, London, 1965, volume 4*, edited by I. Lakatos and A. Musgrave (London and New York: Cambridge University Press, 1970), 51-58.

"Replies to My Critics," *The Philosophy of Karl Popper, Volume 2*, edited by P. A. Schilpp (La Salle: Open Court, 1974): 962-1197.

"A Note on Verisimilitude," *British Journal for the Philosophy of Science* 27 (1976): 147-159.

"Natural Selection and the Emergence of Mind," *Dialectica* 32 (1978): 339-355.

Bibliography

T. E. Hansen, "Bibliography of the writings of Karl Popper" [1925-1974], in *The Philosophy of Karl Popper, Book II*, edited by P. A. Schilpp (La Salle: Open Court, 1974), 1201-1287.

Biographies

M. H. Hacothen, *Karl Popper – The Formative Years, 1902-1945: Politics and Philosophy in Interwar Vienna*, (New York: Cambridge University Press, 2000; Cambridge: Cambridge University Press, 2001).

D. Miller, "Sir Karl Raimund Popper, C. H., F. B. A.," *Biographical Memoirs of Fellows of the Royal Society* 43 (1997): 367-409.

J. Watkins, "Karl Popper: A Memoir," *The American Scholar* 66 (1977): 205-19.

J. Watkins, "Karl Raimund Popper 1902-1994," *Proceedings of the British Academy* 94 (1997): 645-84.

Other References

T. W. Adorno *et al.*, *The Positivist Dispute in German Sociology*, translated by G. Adey and D. Frisby (Heinemann: London, 1976).

J. Agassi, *A philosopher's apprentice: in Karl Popper's workshop* (Amsterdam and Atlanta: Rodopi, 1993).

R. Bambrough, editor, *Plato, Popper and Politics* (New York: Barnes and Noble, 1967).

W. Berkson and J. Wettersten, *Learning from Error: Karl Popper's Psychology of Learning* (La Salle: Open Court, 1984).

M. Bunge, editor, *The Critical Approach to Science and Philosophy: In Honor of Karl R. Popper* (New York: Free Press of Glencoe, 1964).

G. Curie and A. Musgrave, editors, *Popper and the Human Sciences* (The Hague: Nijoff, 1985).

I. Jarvie and S. Pralong, editors, *Popper's Open Society after fifty years* (London and New York: Routledge, 1999).

- I. Lakatos and A. Musgrave (editors), *Criticism and the Growth of Knowledge: Proceedings of the International Colloquium in the Philosophy of Science, London, 1965, volume 4* (London and New York: Cambridge University Press, 1970).
- P. Levinson, editor, *In pursuit of truth: Essays on the Philosophy of Karl Popper on the Occasion of His 80th Birthday* (Atlantic Highlands, NJ: Humanities, 1982).
- B. Magee, *Popper* (London: Fontana/Collins, 1973).
- D. Miller, "Popper's Qualitative Theory of Verisimilitude," *British Journal for the Philosophy of Science* 25, pp.166- 77.
- Miller, *Critical Rationalism* (Chicago: Open Court, 1994).
- A. O'Hear, editor, *Karl Popper: philosophy and problems* (Cambridge: Cambridge University Press, 1995).
- P. A. Schilpp, editor, *The Philosophy of Karl Popper*, 2 volumes (La Salle: Open Court, 1974).
- J. Shearmur, *The Political Thought of Karl Popper* (London: Routledge, 1996).
- C. Simkin, *Popper's Views on Natural and Social Science* Leiden: Brill, 1993).
- P. Tichy, "On Popper's Definition of Verisimilitude," *British Journal for the Philosophy of Science* 25 (1974): 155-60.

Karl Popper's Papers

The Hoover Institute, Stanford University, is the main repository for Popper's papers. These include manuscripts of Popper's works from 1927 to 1980, correspondence from 1932 to 1987, teaching materials from 1937 to 1970, conference proceedings, letters of reference, and additional biographical material. The register can be accessed via the Karl Popper Web: www.eeng.dcu.ie/~tkpw/. Microfilmed copies of the Hoover originals are held by The British Library of the London School of Economics, the Karl Popper Institute (Vienna), the *Karl-Popper-Sammlung* (University of Klagenfurt, Austria), and The Popper Project (Central European University, Budapest).

Chapter Two

A Critique of Popper's Views on Scientific Method

(First Published in *Philosophy of Science*, vol. 39, no. 2, June 1972, pp. 131-152)

Abstract

This paper considers objections to Popper's views on scientific method. It is argued that criticism of Popper's views, developed by Kuhn, Feyerabend, and Lakatos, are not too damaging, although they do require that Popper's views be modified somewhat. It is argued that a much more serious criticism is that Popper has failed to provide us with any reason for holding that the methodological rules he advocates give us a better hope of realizing the aims of science than any other set of rules. Consequently, Popper cannot adequately explain why we should value scientific theories more than other sorts of theories ; which in turn means that Popper fails to solve adequately his fundamental problem, namely the problem of demarcation. It is suggested that in order to get around this difficulty we need to take the search for explanations as a fundamental aim of science.

I

In this paper my aim is to discuss some of the more serious difficulties that Popper's theory of scientific method runs into.

In recent years a number of criticisms of Popper's views have been developed by Kuhn, Feyerabend, Lakatos, and others. I shall argue that these criticisms, in so far as they are valid, follow a common pattern, and can be seen as special cases of a rather more general criticism of Popper. I shall argue that this general type of criticism of Popper is not in fact too damaging, although it does require that Popper's views be modified somewhat.

There is, however, a rather different criticism to be made of Popper's theory, which is much more serious. It amounts to the claim that Popper has failed to provide a rationale for the methodological rules he advocates, and has thus failed to provide an adequate solution to his fundamental problem — namely the problem of demarcation.

Before going any further, I should like to say that, despite my criticisms, I have the greatest admiration for Popper's writings on scientific method (Popper, 1957, 1959 and 1963). I am in complete sympathy with his basic intentions and values. My hope in criticizing Popper's theory is to clarify the problems which confront the theory so that we may have a clearer idea of how the theory can be strengthened and improved so as to cope with these problems.

II

The heart of Popperian methodology may be expressed like this. As scientists, in our hopeful search for the Truth, in our attempt to solve problems of ever increasing profundity, we put forward wild, improbable conjectures, of ever increasing empirical content and explanatory power, which we then seek to overthrow by subjecting them to as severe experimental testing as possible. At any given stage the best theory is the theory of highest empirical content which has stood up best to all our attempts at experimental refutation.

The methodological rules advocated by Popper—governing the acceptance and rejection of theories in science—are designed to give us the best hope of realizing the above aims. According to Popper a supreme rule may be laid down governing the choice of other methodological rules. It is the rule "which says that the other rules of scientific procedure must be designed in such a way that they do not

protect any statement in science against falsification" (Popper, 1959, p. 54).

One point—about which Popper himself has been a bit equivocal—must be made straight away. Scientific theories cannot be refuted experimentally with absolute certainty. This is due in part to Duhem's point that it must always be a group of theories that is tested experimentally, individual theories strictly being incapable of being tested (see Duhem, 1962, ch. VI). But more generally, it is due to the fact that in order to refute with certainty any scientific theory we must establish with certainty the truth of some falsifying hypothesis—and, this, we may take it, cannot be done. It is thus always a *conjecture* that a theory, which our methodological rules leads us to reject, has in fact been found to be false. There is, in other words, always the possibility that the application of Popperian rules may lead us to reject a theory which is in fact true. This situation is not, however, too disastrous—as long as it is reasonable to hold that Popperian rules give us a good hope of correctly detecting error in our theories.

Popper has been at pains to emphasize that a theory can only be falsified with respect to the application of methodological rules (see for example Popper, 1959, pp. 81-2). But this formulation of the issue obscures the decisive point—that, for all we know, application of Popperian rules may lead us again and again to reject true theories. Popper speaks as if we *know* somehow that Popperian rules cannot lead us to reject true theories.

I turn now to a consideration of some of the criticisms that have been made of Popper's theory.

In the first place it has sometimes been argued, in effect, that Popperian rules are not, or have not been, followed in actual scientific practice. Popper's rules thus stand refuted by the empirical evidence, and should be rejected.³

This line of argument is just not cogent. For it is only to be expected that, at the very least, some scientists will sometimes follow bad

³ At least a part of Kuhn's case against Popper is of this type: see for example (Lakatos and Musgrave, 1970, pp. 4-7). Pp. 237-238 also show Kuhn's commitment to the validity alleged of this type of argument. Elsewhere in this book the same type of argument is employed; e.g. by L. Pearce Williams, p. 50, and by Lakatos, p. 115.

methodological policies. And it is *possible*, despite the apparent enormous success of the empirical sciences, that most scientists have most of the time followed not the best of methodological policies. Thus the fact that a theory of method does not "square" with scientific practice does not in itself constitute a refutation of the theory—although of course it may lead us to suspect that there is something wrong with the theory.

A methodological theory simply does not assert that scientists *in fact* follow such and such methodological rules of appraisal. Rather a methodological theory should, I suggest, specify (a) a fundamental aim, or group of aims, for science, and (b) a set of methodological rules; it should then assert:

- (1) The specified aim is the most worthwhile aim for science that is, as far as we know, in principle realizable.
- (2) The specified methodological rules give us the best hope of realizing the specified aim.

In criticizing a theory of method it is *these* assertions, (1) and (2), that need to be criticized. The fact that the specified methodological rules are not followed in practice does not *in itself* undermine either (1) or (2).

One slight qualification needs to be added to this. A theory of scientific method, in order to be acceptable, must have at least *some* contact with scientific practice, with the aims and appraisals of working scientists. A theory of method that has no such contact whatsoever—even though it is perfectly satisfactory in all other respects—could be dismissed out of hand as being irrelevant to even the most liberal idea of what constitutes scientific enquiry.

I do not think however that anyone would want to dismiss Popper's theory in this cavalier fashion.

It might be asked: Suppose an internally consistent methodological theory is developed which is closer to scientific practice than Popper's theory is. Should we automatically prefer the rival theory to Popper's theory?

The answer is no. The rival theory may be based on an aim for science which we may consider to be not so worthwhile as the aim for science proposed by Popper.⁴

A failure, then, of Popper's theory of method to reflect accurately scientific practice does not in itself amount to a flaw in that theory: we may, however, take it as a hint that there may be something wrong with the theory. Here we are making the working assumption (which may well be false after all) that scientists in practice mostly adopt the very best methodological policies.

An entirely different, and rather more valid, type of criticism of Popper's theory, that has been developed by Feyerabend (1970), Lakatos (1970), and at least by implication by Kuhn (1970a), amounts to the claim that Popper's methodological rules do not give one the best hope of realizing Popper's aims for science. Situations arise, it is argued, in which following Popper's methodological rules would seriously impede the growth of science.

Almost all of Popper's methodological rules have been criticized in this way. Here are in turn some main Popperian methodological rules (see Popper, 1959) and, very briefly, the criticisms which have been made of them.

- (1) An acceptable new theory must always have greater empirical content than its predecessors.

This is too severe. In certain circumstances it will be in our interests to accept a new theory which initially has far less empirical content than its predecessors, precisely because the new theory promises, with development, with the addition perhaps of auxiliary hypotheses,

⁴ Of course a methodologist may be concerned to characterize not, as Popper is, the highest ideal of science, but rather science as it in fact exists. In this case the methodologist will select those fundamental aims for science which are such that the rules best adapted to realize these aims are as close as possible to rules in fact adopted by the majority of scientists. In this case of course the "correct" methodological theory will be that internally consistent theory which is closest to scientific practice.

This, on a charitable interpretation, is perhaps Kuhn's conception of methodology.

to lead to a theory of far greater empirical content than its predecessors.⁵ The new theory may for example solve outstanding theoretical problems that the old theories were unable to solve.

It is in any case important to develop rival theories to the existing theories, even if these rivals have nothing like the empirical content of the existing theories, for often it is only by developing such rival theories that we can test the existing theories.⁶

- (2) An acceptable new theory must at least be able to explain all of the past success of its predecessors.

Even this more modest demand is too severe. A new theory may be acceptable even though it cannot explain much that its predecessors could explain, for again the new theory may solve severe outstanding problems, and thus promise to lead to a theory which in the end explains all that the old theories explained, and much more besides.

In his later writings Popper has stressed the importance of seeing scientific enquiry as a problem solving activity. Neither Popper nor any of his critics has, however, realized that the demand that a new acceptable theory should solve outstanding problems may conflict with and, on the short-term, actually override the demand that a new acceptable theory should have excess content over its predecessors.

- (3) Always test a theory as severely as possible.

This assumes that in testing a theory our invariable concern is to seek to falsify it. But this assumption is false. In testing a new theory in particular our concern may be to develop the theory, extend the range of its successful applications, build up auxiliary hypotheses. And in order to do this it may well be in our interests to test initially only the most straightforward, least problematical implications of the new theory. We may be justified in actually ignoring, for a time, refuting instances of a theory. For even if we have good grounds for suspecting a

⁵ See Feyerabend (1970) for a powerful development of this criticism. Feyerabend, however—in my view quite wrongly—regards this criticism as spelling the downfall of Popper's theory of scientific method.

⁶ This point is made by Feyerabend in his (1965). One might add that Popper's notions of severe testing, and of corroboration, actually require that when we test a theory we have some rival theory up our sleeve.

theory to be false, it may be in our interests to develop the theory further, as this may indicate more clearly what a new theory must ultimately explain. There is, in short, a methodological point to what Kuhn has called "normal" science. "Revolutions" in science may only be profitable after a phase of normal science (Kuhn, 1970a).

In short, our concern ultimately is to test severely all our theories. But it is theories that have been allowed to develop, grow and reach their full strength that we wish to test severely. It may well be against our interests to test severely, and perhaps eliminate, a promising theory which has not been allowed to come to full fruition.

There is another point. As Lakatos (1970) has pointed out, in a research programme, refutations may be completely unsurprising and in a way expected, because, to begin with, simplifying assumptions, known to be false, are made. It may be not the refutations, but rather the corroborations, which are unexpected and surprising. Lakatos considers the following assumptions made by the Newtonian programme to explain in detail the motions of the planets.

- (i) Masses of heavenly bodies are concentrated at their centres in mass points of infinite density.
- (ii) The sun is stationary.
- (iii) The planets do not interact gravitationally.
- (iv) There are no tidal effects.

These assumptions were successively dropped as the programme became more and more sophisticated. The point to note is that each of these assumptions is actually inconsistent with Newton's laws of motion plus the law of gravitation. If the laws are true, then these simplifying assumptions must be false. In these circumstances there is in a sense no point in testing severely an early crude application of Newtonian theory to the solar system. What is needed rather is the gradual development of the Newtonian programme (in this case the development needed was largely mathematical) so that eventually a sophisticated version of Newtonian theory *can* be severely tested, a version that does not incorporate inconsistent assumptions.

- (4) An experimentally "refuted" theory must be rejected.

As before this is too drastic. In general it will not be in our interests to reject a theory that, in the past, has had considerable empirical

success until there is an alternative more promising theory on the horizon: see Lakatos (1970).

(5) An experimentally "refuted" and rejected theory must not be revived at a later stage.

A true theory (or at least an extremely valuable, promising theory) may be "refuted" and quite properly rejected, the "refutation" being due to false auxiliary hypotheses. In order to take this possibility into account, the Popperian rule must be rejected. We can, however, say this: we cannot simply return to an earlier theoretical situation. A theory that has been discarded can only be reaccepted if auxiliary hypotheses, that have been developed subsequently, considerably increase the empirical content of the theory.

(6) An inconsistent theory cannot be accepted.

It may well be in our interests to accept provisionally a formally inconsistent theory, in the hope that further theoretical work will remove the inconsistency.

It may be asked : Granted that all these criticisms are cogent, do they not effectively demolish Popper's theory? Instead of trying to patch up Popper's theory in the light of these criticisms, should we not reject the theory altogether, and try to find some new and more adequate methodological theory ?

As I see it, the situation is like this. Each of the above objections, (1)-(6), is in effect a particular illustration of the general point that if we wish to follow Popper's rules in the long run, then, in certain circumstances, we will be well advised to break these rules on a short term basis. Each criticism argues in effect that it is against our interests to enforce too rigidly Popper's essentially long-term strategic rules on the short-term, tactical level.

But this sort of situation is almost bound to occur given almost any complex, long-term, goal-directed activity. Long-term and short-term interests are almost bound to clash at times. It is to be expected that strategic interests will at times make desirable modes of action which, if judged on a purely short-term, tactical level, would seem to be highly undesirable. Or, in other words, given almost any complex goal-directed activity, it is to be expected that it will be against our interests to enforce too rigidly strategic rules on the tactical level.

It is thus no inadequacy whatsoever in Popper's strategic rules that they too suffer from this quite general complaint. It would in fact be quite extraordinary if they did not.

The above criticisms, (1)-(6), do not effectively demolish Popper's position. They simply spell out in detail the general point that short-term exceptions can invariably be found to long-term methodological rules. It is precisely this general point that Popper has failed to recognize, and it is this failure which vitiates the presentation of his theory.

It is not altogether surprising that Popper has ignored this point. In his (1959), a major concern of Popper is to solve the problem of demarcation. Consequently, Popper is in the main concerned with those large scale, strategic methodological rules in terms of which (a) scientific enquiry can be distinguished from other types of enquiry; (b) criteria of overall scientific growth can be formulated. Popper, quite understandably, does not stop to consider the detailed application of his methodological rules on the "tactical" level.

Popper's position, then, needs to be modified in the light of the criticisms (1)-(6).⁷ But this modification does not amount to an ad hoc patching up of Popper's theory. Rather the modification is one which any first formulation of a methodological theory would have to undergo, whatever the game might be — science, war, chess, economic planning, or whatever.

III

I come now to my own major criticism of Popper's theory. It amounts to this: Popper has failed completely to provide any kind of rationale for the methodological rules he advocates. That is, he has failed to provide us with any reason for holding that Popperian rules give us a better hope of realizing the aims of scientific enquiry than any other set of rules. Nor is it easy to see how this failure can be made good within a general Popperian framework. Consequently, Popper has failed to solve his fundamental problem—the problem of demarcation. He has also failed to exhibit science as a rational enterprise. For in

⁷ Such a modified version of Popper's theory has been put forward by Lakatos (1970).

order to do this, it is not enough simply to specify an aim for science and a set of methodological rules: we need in addition some reason for holding that the rules give us a better hope of realizing the aim than any other set of rules.

In order to develop this criticism, I want to begin by considering—how Popper himself tackles his fundamental problems—which I take to be the problems of demarcation and induction.

The traditional problem of induction, as considered by Popper, can be put like this: How is it possible, by verifying particular instances of a strictly universal statement, or scientific law, to verify the universal statement or law itself, with at least some degree of probability greater than zero? The problem of demarcation, again as considered by Popper, can be put like this: What is the distinctive feature of scientific theories which enables us to distinguish scientific theories from other sorts of theories? ⁸

⁸ Popper here conflates three distinct demarcation problems, namely the problems of distinguishing:

(i) Empirical theories from other sorts of theories (metaphysical, a priori, etc.).

(ii) Scientific modes of enquiry from unscientific or pseudoscientific modes of enquiry,

(iii) The highest form of scientific enquiry from other less worthwhile forms of scientific enquiry grading into the pseudoscientific.

A few remarks about these three problems.

Popper clearly is primarily concerned to solve the third of these three demarcation problems. Kuhn, on the other hand, I have suggested, may be interpreted as being concerned to solve the second demarcation problem. And the logical positivists sought, and failed to discover, a solution to the first demarcation problem, in so far as their principle of verification was intended to be a criterion of empiricalness, and not, absurdly, a criterion of meaning.

It seems to me to be desirable to regard 'scientific' as being attributable, in the first instance, to modes of enquiry, rather than, as Popper would have it, to theories, or even as Lakatos would have it, to series of theories. What is most characteristic of science is its methods of investigation, rather than the type of theory it deals with. We can of course then go on to apply 'scientific' to a theory which is proposed as part of the course of scientific investigation.

With respect to these two problems, Popper's position, put crudely, amounts to this: the distinctive and especially valuable feature of

It also seems to me to be desirable to distinguish an empirical theory from a scientific theory. A theory may be empirical in Popper's sense of being experimentally falsifiable, and yet may not be proposed in the context of a scientific enquiry. We need to be able to say that a mode of enquiry considers only empirical theories, and yet is not scientific: Popper's terminology does not permit this. [Note added in 2010. Not true: a mode of enquiry might only consider empirical theories and yet fail to be scientific, as this is construed by Popper, because successive theories have decreasing empirical content.]

It should be noted, however, that a theory can only be said to be empirical or experimentally falsifiable relative to the acceptance of at least two distinct sorts of methodological rules. First, there must be agreement about what sort of statement can constitute a basic statement. Second, there must be agreement about under what circumstances a falsifying hypothesis is to be accepted. These are not necessarily trivial matters. In psychology, for example, there is no general agreement about what sort of statement qualifies as a basic statement.

It may be asked: If only conglomerations of theories are experimentally falsifiable, and not individual theories, how can we define the empirical character of an individual theory in terms of falsifiability? This can, I think, be done (using the concept of empirical content) along the following lines. A hypothesis h is empirical if, and only if:

(a) There is a falsifiable conglomeration of hypotheses, T , which entails h ; and which is such that:

(b) There is no T' , with the same empirical content as T , such that T entails T' , but T' does not entail h .

(In other words we cannot omit h from T without decreasing the empirical content of T .)

This seems to correspond to our intuitions. For example we are inclined to call Newton's postulate concerning absolute space metaphysical, i.e. non-empirical, precisely because we can remove this postulate from Newtonian theory without thereby decreasing the empirical content of that theory.

Here then I suggest is a solution to that age-old problem of giving a precise formulation to the so-called principle of verification. Note: I do not wish to suggest that metaphysical components of a scientific theory are necessarily undesirable.

scientific theories is that they are experimentally falsifiable. Falsifiable theories are especially to be prized just because we can discover that they are false; in this field we can detect error, learn from our mistakes, and so hopefully make progress. Scientific theories cannot be verified, or even rendered probable, by any amount of experimental evidence: the traditional problem of induction, in other words, is insoluble. But this does not matter, as long as we give a non-inductivist, non-justificationalist solution to the problem of demarcation. It is only if we hold that the distinctive and especially valuable feature of scientific theories is that they are amenable to inductive *verification* that it becomes an urgent matter to solve the problem of induction.

In short, Popper offers a solution to the problem of demarcation which, he maintains, makes it unimportant that the traditional problem of induction is insoluble.⁹ Now the point that needs to be emphasized is this. A major part of the tremendous power and suggestiveness of Popper's proposal is due to the fact that Popper's demarcation

⁹ Of course, a rather more general version of the problem of induction may be formulated, which is less obviously insoluble, such as: What criteria ought to govern our choice of a theory from two or more rival theories? (Here we simply drop the implicit assumption that it is high probability which ought to govern our choice of theories.) It at once becomes clear that there are at least two distinct versions of this formulation of the problem of induction, namely:

- (i) What criteria ought to govern our choice of a theory from two or more theories *if our concern is with scientific growth?*
- (ii) What criteria ought to govern our choice of a theory from two or more rival theories, *if our concern is with the trustworthiness of the theory, for purposes of technological application?*

Now Popper does claim, in effect, to have solved the first version of this problem. But about the second version of the problem Popper is altogether silent. He cannot claim to have (a) solved the problem; (b) shown the problem is insoluble; or (c) shown that the problem is unimportant. It might of course turn out that a satisfactory solution of (i) also provides a satisfactory solution of (ii): but this is not obvious. It needs to be argued. (Lakatos in effect sets out to provide something like such an argument in his 1968a.) Popper, however, provides scarcely even a hint of such an argument—apart from the odd stray remark (e.g. 1963, p. 51). Here then is an inadequacy in Popper's views in addition to the one discussed in the text.

criterion enables us to explain why we should value scientific theories *even though they remain utterly improbable conjectures*. The Popperian revolution is, in other words, to a considerable extent, a revolution in *values*. Before Popper scientific theories were valued for their supposed relative certainty, security, high probability. After Popper, scientific theories are valued for their amazing content and explanatory power, and for their vulnerability to one of the most devastating forms of criticism yet to be discovered—experimental testing. And those scientific theories that we continue to accept are valued for their ability to survive all our most searching attempts to overthrow them.

It is clear then from his solution that Popper interprets the problem of demarcation as the problem of providing a criterion of "scientificness" which enables us to explain why we especially prize scientific theories. *A criterion of demarcation which gave no hint whatsoever of why we should value scientific theories more than other sorts of factual theories would be an utterly unacceptable solution to the demarcation problem.*

The demarcation problem that Popper sets out to solve—but does not quite succeed in formulating—can, then, be put like this: What is the distinctive feature of scientific theories which (a) enables us to distinguish scientific theories from other sorts of theories, (b) leads us especially to prize scientific theories?

We are now in a position to consider the fundamental problem: Can we provide a rationale for Popper's revised methodological rules? Do we have any reason for holding that Popper's revised methodological rules give us a better hope of realizing the fundamental aim of science than any other set of rules?

In order to discuss this question we need to specify precisely what we take to be the fundamental aim of science.

Four "Popperian" aims for scientific enquiry may be distinguished, graded here in order of ambitiousness.

- (1) Successively to put forward, and reject, theories in accordance with the methodological rules of acceptance and rejection specified by Popper.

- (2) To put forward, and in turn *falsify*, theories of ever increasing content, which explain all of the empirical success of their predecessors.
- (3) To put forward theories which, in addition to (2), successfully predict phenomena not known at the time of their formulation, before they are eventually refuted.
- (4) To put forward theories of ever increasing degrees of verisimilitude.

Now if aim (1) is accepted, Popper's methodological rules follow *analytically* from the conception of science implicit in (1). Anyone who does not adopt Popper's methodological rules is simply not playing the scientific game as defined by (1). In this case, of course, there is no need to provide a *rationale* for the methodological rules in question. But a high price is paid: this line of approach cannot provide an adequate solution to the problem of demarcation. For, as we have seen, in order to solve the problem of demarcation it is not enough simply to specify necessary and sufficient conditions for a type of enquiry to be scientific: in addition one needs to show why we are justified in especially prizing the theories of a mode of enquiry that is scientific in the required sense. But clearly, the mere fact that a mode of enquiry proceeds in accordance with Popper's acceptance and rejection rules provides no reason whatsoever for especially valuing its theories.

If on the other hand aim (2) is accepted, a provisional reason can be given for especially prizing scientific theories. For one can say of an unrefuted scientific theory that despite the fact that it is vulnerable to refutation, nevertheless it has survived our most earnest attempts to refute it, and thus, *prima facie* at least, deserves to be taken more seriously than any theory which has not been subjected to this particularly devastating form of scrutiny. But at once there arises the problem of providing a *rationale* for our methodological rules, for we can ask the question: Do our rules give us the best hope of testing our theories in the severest way possible? Do our rules give us the best hope of rejecting a theory only when we have isolated some false implication of the theory?

In *The Logic of Scientific Discovery* Popper seems to adopt the view that methodological rules simply define the game of science, and require

no kind of rationale.¹⁰ But then he must accept (1) as his aim for science, and he cannot claim to have solved the problem of demarcation. Clearly Popper would not wish to accept either of these things. He must therefore face the problem of providing a rationale for methodological rules.

In fact of course Popper does not wish to defend (2) as a worthy aim for science either (see his 1963, ch. 10). Taking (3) or (4) as one's aim for science has the advantage of enabling one to give even stronger potential solutions to the problem of demarcation. Thus if one adopts aim (4), one can say that scientific theories constitute our best efforts so far at discovering the truth. But the penalty for this is that the problem of providing a rationale for our adopted methodological rules becomes all the more severe. For we need to establish that our methodological rules give us a better hope of maximizing the verisimilitude of our theories than any other set of rules.

To put it bluntly: because Popper has failed to provide any kind of rationale for the methodological rules he advocates, he has failed to give an adequate solution to the problem of demarcation, and to that version of the problem of induction which he would wish to claim he has solved, namely: What criteria ought to govern our selection of theories if our concern is to realize the fundamental aim of scientific enquiry? In addition he has failed to show that scientific enquiry can be viewed as a rational enterprise. In order to solve these three problems adequately it is essential to show that the advocated methodological rules give one a better hope of realizing the fundamental aim for science than any alternative methodological rules; and it is just this which Popper has failed to do.¹¹

¹⁰ Popper says for example: "Just as chess might be defined by the rules proper to it, so empirical science may be defined by means of its methodological rules," see his (1959, p. 54).

¹¹ It is a certain ambiguity in Popper's notion of "falsifiability" or "refutability" which is in a sense responsible for this failure. Popper holds in effect that scientific theories are not decisively falsifiable—in order to meet the Duhem argument, or "conventionalist objections"—and at the same time holds that scientific theories are decisively falsifiable—in order to hold at bay the rationale problem, and thus ensure that he can give an adequate solution to the problem of demarcation.

In order to drive this criticism of Popperian methodology home, I need to show in greater detail that no rationale for Popper's methodological rules can be provided within Popper's theory, granted that we take in turn (2), (3), and (4) above as the aim for science. Before I do this, however, I would like now to point out that my criticism of Popper is *not* simply a reiteration of the well-known criticism, outlined by for example Salmon (1968), and Lakatos (1968a), that Popper fails to show that highly corroborated theories are also the most reliable or trustworthy for technological applications.

Let us consider Lakatos' exposition of this criticism. According to Lakatos, in addition to standard Popperian methodological appraisals (which Lakatos calls *acceptability*₁ and *acceptability*₂) we need an additional appraisal — *acceptability*₃ — which is a measure of the "'inductive acceptability', 'trustworthiness', 'reliability', 'evidential support', 'credibility', etc." of a theory (1968a, p. 391). We need this additional appraisal solely in order to assess the worth of a theory for, practical, technological purposes. Lakatos remarks: "In formulating Popper's methodology there is no need to refer to *acceptability*₃" (1968a, p. 392).

Now in his (1968a), Lakatos' criticism of Popper amounts to this: the problem of the *acceptability*₃ of theories is both an important problem, and a problem which Popper fails to solve. With this I entirely agree. (It is just this point that I made in footnote 7 above.) This argument of Lakatos is, however, quite different from the one that I am developing in the present section. For my argument here is this: quite aside from Popper's failure to solve the *acceptability*₃ problem (which—as Lakatos notes—Popper regards as "comparatively unimportant"), Popper fails, much more seriously, in his primary purpose; namely, to solve the problem of demarcation. A major concern of Popper is to exhibit science as *rational* enquiry (leaving aside the additional issue of whether science can be shown to be a basis for rational action); and it is just this, I wish to argue, which Popper fails to do; for, he fails to provide a rationale for the methodological rules he advocates. Thus Lakatos criticises Popper for failing to solve a problem which Popper did not really set out to solve in the first place; whereas I wish to criticize Popper's methodology for failing to achieve its primary objective, namely to exhibit scientific enquiry as more

rational than other types of enquiry and thus provide a solution to the problem of demarcation.

I want now to discuss in more detail the problem which confronts us in attempting to provide a rationale for Popperian methodological rules, given that our aim for science is in turn (2), (3), and (4) above.

Let us, to begin with, take (2) as our aim for science; that is, let us assume our aim is:

- (2) To put forward, and in turn *falsify*, theories of ever increasing content, which explain all of the empirical success of their predecessors.

Can we show that suitably revised Popperian rules give us the best hope of rejecting a theory only when we have detected some error in that theory?

In view of the Duhemian argument, we cannot hold that an individual theory can be decisively falsified experimentally. It might be, however, that Popperian rules give us the best hope of detecting error in our theories.

Consider the following situation. There are two rival research programmes, centered round the two conflicting theories T_1 and T_2 . The first research programme, based on T_1 , has for a long time stagnated; a host of auxiliary hypotheses have been proposed to salvage T_1 from rejection; despite this, a great number of well-corroborated hypotheses conflict with T_1 plus auxiliary hypotheses. In addition this research programme has come up with no striking new predictions which have been corroborated.

In contrast to this, the research programme based on T_2 goes buoyantly ahead. The empirical content of T_2 far exceeds that of T_1 . T_2 has made a number of striking new predictions which have been corroborated.

In these highly exaggerated circumstances it is clear that Popperian rules would oblige us to accept T_2 and reject T_1 .

Do we have any grounds, however, within Popper's general viewpoint, for maintaining that T_1 is false? Do we have any rational grounds for preferring the conjecture that we have detected error in T_1 to the conjecture that we have detected error in T_2 ?

We might argue: T_1 is, at least, in difficulties; whereas T_2 is not. Hence it is more rational to suppose T_1 is false.

But suppose T_1 is true—something which is perfectly possible. Suppose further that we have adopted some non-Popperian methodology which obliges us to accept T_1 and reject T_2 , in the above circumstances. In this case it would be T_2 that would be in very severe difficulties, since it would be "refuted" relative to our acceptance of T_1 . In these circumstances, adoption of the non-Popperian methodology would lead us to detect genuine error in T_2 , whereas adoption of the Popperian methodology would lead us to suppress the refutation of T_2 .

We might argue: the research programme based on T_1 has stagnated, whereas the research programme based on T_2 forges ahead. It is more rational in these circumstances to suppose that T_1 is false, rather than T_2 .

But why? Perhaps the universe is so constructed that those research programmes which forge brilliantly ahead are precisely those which plunge us deeper and deeper into error. It might be some utterly stagnated programme which contained the seeds of truth, and which, after further great effort, might enable us to glean a little more of the truth. Popper provides us with no rational grounds for excluding such a possibility.

Finally we might argue: in the above circumstances T_2 has been more highly corroborated than T_1 ; T_2 thus gives us a firmer assurance of being true than T_1 .

But no Popperian can employ this argument, since to do so is to resort to the despised inductive approach to scientific method.

I conclude that, within the general Popperian standpoint, no reason can be given for holding that Popperian rules give us the best hope of rejecting individual theories only when we have detected error in them.

Individual theories cannot be decisively refuted; suitable conglomerations of theories—so it may be argued—can, however, be decisively refuted. It may be held, then, that Popperian rules give us the best hope of detecting error in conglomerations of theories.

But even this more modest claim cannot it seems be maintained. For according to Popper a suitable conglomeration of, let us say, relatively high-level theories is only refuted relative to the acceptance of low-level, experimental, falsifying hypotheses. And these falsifying hypotheses, however well-corroborated, cannot, according to Popper, be established

with any more certainty than the high-level theories. Thus we do not have decisive refutations of even conglomerations of theories.

We do not even have any reason for supposing that Popperian rules give us the best hope of detecting error in conglomeration of theories. For suppose an enormous number of extremely well-corroborated, low-level, experimental hypotheses conflict with some conglomeration of theories. In this case, if all hypotheses are equally improbable, why not regard the conglomeration of theories as *refuting* the experimental hypotheses, instead of the other way round? If all that we have is two sets of hypotheses that contradict each other, we can have no reason for preferring the conjecture that error has been detected in one set rather than in the other set.

In reply to this a Popperian might argue that a falsifying hypothesis may be a singular existential statement, which can be established with more security than any universal statement. Thus refutations, if not exactly decisive, are nevertheless reasonably assured by Popperian rules. In order to refute 'All ravens are black' for example we only need to establish: 'There is a family of white ravens in the zoo at New York' (Popper's own example, see 1959, p. 87). This falsifying hypothesis is intersubjectively testable; it describes what is in a sense a repeatable effect—since the hypothesis does not simply assert 'There is a white raven at such and such a time and place'. Nevertheless the hypothesis is a singular statement.¹²

¹² One reply to this Popperian counterproposal is to argue, as Deutscher (1968) in effect does, that Popper provides us with no reason for supposing that the methodological rules that govern acceptance of singular existential statements are the best available if we wish only to accept true basic statements. Deutscher is, I think, right in stressing that Popper fails to solve this problem adequately. However, it seems reasonable to suppose that a solution can be found to this problem within Popper's general methodological position. But a solution to this problem would not really help Popper at all. For, as I argue in the text, acceptance of singular existential statements cannot lead to the overthrow of physical theories. It thus, within the Popperian framework, remains utterly problematic how we could ever have good grounds for supposing a theory to be falsified. I do not think, therefore, that Deutscher has isolated the most severe part of the problem that confronts Popper's theory.

There are a number of objections to this line of argument. Consider the law 'All pieces of copper expand when heated', and suppose a particular piece of copper is discovered which fails to expand when heated. (For simplicity we consider an individual law rather than a conglomeration of laws.)

Now in the first place the statement 'This piece of copper fails to expand when heated' would appear to incorporate universality to as great a degree as any rather more straightforward more universal statement. For the statement asserts, in effect, 'At all times, and in all places, this piece of copper fails to expand when heated'. Hence this statement would appear to be, on Popperian grounds, as impossible to verify as any straightforward universal statement.

To this it may be retorted: but the falsifying hypothesis can quite easily be transformed into a genuinely singular statement: e.g. it could be formulated as 'During the next year (or ten minutes), anywhere on the earth's surface (or anywhere in this laboratory) this piece of copper will fail to expand when heated'.

One might query whether even this severely circumscribed falsifying hypothesis can be verified with any more certainty than any universal statement, since the hypothesis still carries implications about an indefinite number of experiments. But let that pass.¹³ The important point to notice is that, given this last formulation, the hypothesis would not in fact be sufficient to overthrow the physical law 'All bits of copper expand when heated'. For in accepting experimental results as refuting a theory, one is committed to the possibility of explaining these results by some future theory. That is, one is committed to holding, at least as a conjecture, that the refuting experimental results constitute lawful occurrences. For if one denies this conjecture one thereby accepts experimental results which no future physical theory can conceivably

¹³ Popper could not in fact let this pass, for he holds that even a singular existential statement will involve some law, in that the statement will attribute a dispositional property to some object (see his 1959, pp. 423-424). I am not so sure, however, that in attributing a dispositional property to an object we are necessarily, if implicitly, asserting some law-statement. It seems to me to be conceivable that genuine singular existential statements, not involving laws, but nevertheless specifying a repeatable, intersubjectively testable effect, are possible.

explain—since it is only lawful occurrences that can be explained physically. Clearly, physics will be highly resistant to accepting the existence of such phenomena. But to assert that a set of experimental results constitute lawful occurrences is in effect to assert a somewhat vague *universal* hypothesis. Thus *singular* hypotheses, however well-corroborated, do not suffice to refute a theory; it is only singular hypotheses, backed up by a *universal* hypothesis, to the effect that the experimental results in question constitute lawful occurrences, that can refute a theory—or rather a conglomeration of theories.

Consider again the law 'All bits of copper expand when heated'. Let us suppose 'This piece of copper, on the surface of the earth, for one year, fails to expand when heated' is highly corroborated within the specified limits of space and time, but that outside these limits the piece of copper in question expands normally when heated. Suppose also that, apart from the above exception, 'All bits of copper expand when heated' continues to be corroborated.

Given these somewhat extraordinary circumstances, we might well decide to retain the law 'All bits of copper expand when heated', and simply declare that in the case of the bit of copper which, for a time, failed to expand when heated, something was going on which we do not understand, and cannot as physicists take into account. That this might be a legitimate decision, in these circumstances, indicates that the singular statement 'This piece of copper, for a time, fails to expand when heated' cannot, however well-corroborated, in itself conclusively refute 'All bits of copper expand when heated'.

However, in the above circumstances we should not I think be obliged to take this line. For we could always conjecture that the anomalous bit of copper was, during the relevant time, in an unusual state, or exposed to unusual conditions, and in the light of this conjecture, reject 'All bits of copper expand when heated'. But the crucial point to note here is that we should be rejecting 'All bits of copper expand when heated', not as a consequence simply of accepting the singular hypothesis 'This bit of copper, for a time, fails to expand when heated', but rather as a consequence of accepting the somewhat vague, *universal* hypothesis 'All bits of copper, when in some specific unknown state, fail to expand when heated'. It is only if we accept tentatively some such hypothesis as this that the behavior of the anomalous bit of copper

can be seen as constituting lawful occurrences, which we may hope to explain by means of some future theory.

Of course 'All bits of copper, when in some specific unknown state, fail to expand when heated' is, in Popper's terminology, a *metaphysical* hypothesis. But this does not affect the argument. Acceptance of this "metaphysical" hypothesis does not mean that the hypothesis is accepted as a full-fledged respectable scientific theory; rather, it means that one is committed to a certain research programme, namely to develop a precise, non-metaphysical version of the vague metaphysical hypothesis. If this research programme is wholly unsuccessful then we must reject the ostensible refutation of the law 'All bits of copper expand when heated', especially if acceptance of this law leads to a highly successful research programme. All of which shows that 'This piece of copper, at such and such times and places, fails to expand when heated' however well-corroborated, cannot on its own, conclusively refute 'All bits of copper expand when heated'.

I conclude that no conglomeration of theories can be conclusively refuted, and more generally, that we have no grounds for maintaining within the general Popperian viewpoint, that Popper's methodological rules give us the best hope of detecting error in either individual theories, or in any conglomeration of theories.¹⁴

As I have already remarked, Popper does not accept that it is sufficient for science to aim at putting forward theories of ever increasing content which are in turn refuted (see his 1963, ch. 10). He holds that science must in addition aim at putting forward theories which have a certain measure of empirical success before they are eventually refuted. If we drop the aim of refutation, we may consider the following somewhat revised Popperian aim for science, a revised version of aim (3) above:

- (3)' To put forward theories of ever increasing content, which (a) explain the empirical success of their predecessors (b) make

¹⁴ In order to avoid a possible misunderstanding, I should perhaps add that I believe that something like Popperian rules do give us the best hope of detecting error in our theories. The problem however is to show that we have good reasons for this thesis—especially within Popper's general methodology.

successful new predictions before running into severe empirical difficulties.

Can we maintain that revised Popperian rules, or let us say, Lakatosian rules, give us the best hope of realizing this aim for scientific enquiry? This is not an easy question to discuss, because of a crucial ambiguity in the aim specified in (3)'. Terms such as 'refutation', 'corroboration', 'empirical success', 'successful prediction', and 'empirical difficulties' can be interpreted in two radically different ways. On the one hand these terms may be explicated solely in terms of certain adopted methodological rules, no attempt being made to provide a rationale for these rules. In this case 'refutation', for example, simply means that the methodological rules require that we reject the theory in question. Again, 'successful prediction' means simply that the theory in question predicts a hypothesis which our rules require us to accept. We may call this the "methodological" interpretation of the above terms.

On the other hand the above terms may be given what we might call an "epistemological" interpretation. Here the assumption is that we have specified some aim for science which enables us to explain why we especially prize scientific theories, and hence which enables us to give an adequate solution to the problem of demarcation. 'Refutation', 'successful prediction', etc., are then explicated in terms of this aim. Thus the epistemological interpretation of 'refutation' might be that we have good, or rational, grounds for holding the theory in question to be false. The epistemological interpretation of 'successful prediction' might be that the prediction should be taken seriously in that it has survived our very real attempts to falsify it.

Now if the relevant terms in (3)' above are given a methodological interpretation only, then it is clear that the aim specified in (3)' cannot constitute a satisfactory aim for science. For such an aim amounts to no more than the aim to accept and reject theories in accordance with Popperian, or Lakatosian, rules of acceptance and rejection. In terms of such an aim we can give no explanation of why we especially prize scientific theories; we cannot provide an adequate solution to the problem of demarcation.

If on the other hand we interpret the relevant terms of (3)' epistemologically, then all the old problems arise afresh. Either we give a non-Popperian, justificational interpretation to the relevant terms, or we give a Popperian interpretation, which commits us to maintaining that Lakatosian rules give us the best hope of detecting error in our theories. And this, we have already argued, cannot be maintained. There just does not seem to be a third interpretation available. If 'refutation' has nothing to do with the desired detection of falsehood, and 'corroboration' has nothing to do with the desired detection of truth, then it seems we have abandoned what I have called the "epistemological" realm altogether. Scientific investigation becomes simply an intricate game, the only purpose of which is to play the game in accordance with the rules.¹⁵

One might of course attempt to give an extra-methodological point to aim (3)' in terms of the claim that highly corroborated theories are more trustworthy, for technological purposes, than uncorroborated or refuted theories. But apart from the intrinsic difficulties associated with this claim—it involves something like the traditional problem of induction—to adopt this line would be to give essentially an instrumentalist defence of the value of science, which would be to break radically with the Popperian tradition.

There is finally Popper's fourth aim for science to consider, namely:

- (4) To put forward theories of ever increasing degrees of verisimilitude.

But this case clearly requires no additional discussion. If one cannot even provide a rationale for Popperian rules given the relatively unambitious aims for science, (2) and (3), then quite clearly there can be no hope of providing a rationale for Popperian rules, within Popper's general viewpoint, given the highly ambitious aim for science to approach closer and closer to the truth.¹⁶

¹⁵ A somewhat similar argument has been developed independently by Lakatos; see Lakatos' contribution to Schilpp (1974).

¹⁶ Of possible relevance to the problem of providing a rationale for Popperian rules given the aim of increasing the truth-content of our theories, is Popper's thesis that the only way to increase truth-content is to increase content (see his 1970, p. 57 and 1966). But this at most

There is, however, one reply that Popper might wish to make to my whole argument so far. Popper might argue that it must always be a *conjecture* that such and such rules give one the best hope of realizing the aim of science—whether that aim is the detection and elimination of error, or the discovery of the truth. One cannot *know for certain* the methodological rules one adopts are the best available. In this domain too one can only put forward conjectures, and then seek to refute them, by subjecting them to severe criticism.

This reply does not however dispose of my argument. On the contrary, my whole argument can be interpreted as being designed precisely to show that, within Popper's general standpoint, there can be no reason for preferring one conjecture to another as to the best rules to adopt if one wishes to detect error, or progress towards the truth. In other words I have argued, in effect, that within Popper's general standpoint, conjectures about which rules it is best to adopt are all equally uncriticizable. It is in this sense that we have no grounds, no reasons, for preferring one conjecture to another.

IV

No rationale can then, it seems, be provided for Popperian methodological rules, given Popper's general methodological viewpoint. It is natural to ask: Can a rationale be provided for Popperian rules which does not do violence to the whole spirit of Popper's philosophy, even though it may conflict with one or other of Popper's rather more detailed methodological theses? I am inclined to think that such a rationale can be provided. But first the terms of the discussion need to be changed a little.

So far the problem has been discussed in the following terms. Granted that the aim of science is to progress towards the truth (or, more modestly, to detect and eliminate error) how can we show that methodological rules roughly similar to those advocated by Popper give us a rational hope of realizing this aim? In my view the problem posed in this way makes the fundamental mistake of supposing that the aim of science is simply to progress towards the truth (at its most

provides a rationale for just the rule: put forward theories of ever increasing empirical content. It does not in any way provide a rationale for all the other Popperian rules.

ambitious). For the fact is surely that a basic aim of science is to search for *explanations*. Science seeks, not truth *per se*, but rather what might be characterized as *explanatory truth*. That is, science seeks to put forward theories of *both* increasing explanatory power *and* increasing verisimilitude.¹⁷ The possibility that the truth is non-explanatory is something which science simply disregards, even though we have no good reason to suppose that the truth is explanatory rather than nonexplanatory.

The notion of "explanatory power" or the closely related methodological notion of "simplicity" is of course notoriously difficult to characterize precisely. Roughly one can say that the more a theory predicts, and the fewer the number of independent postulates it has, so the greater is the theory's degree of simplicity, or of explanatory power. A theory which is so "unsimple" as to have as many independent postulates as the number of laws it entails is not explanatory at all.

Now Popper is, of course, ready to acknowledge that a basic aim of science is to search for explanations (see the example 1957, and 1963, pp. 114-115). However, Popper has repeatedly tried to reduce this aim to an even more fundamental aim. A basic tenet of his (1959) is that simplicity, or explanatory power, is to be equated with empirical content. And Popper has tried to reduce the aim of searching for theories of ever higher empirical content to the aim either of detecting and eliminating error, or of approaching the truth. In his (1959) Popper may be interpreted as taking the detection and elimination of error as the fundamental aim of science. In order to realize this aim, we need to consider theories vulnerable to refutation, and in particular theories of ever increasing degrees of falsifiability (and of course degree of falsifiability = degree of empirical content). Later Popper argues that if we pursue the fundamental aim of approaching the truth we need to consider theories of ever increasing degrees of empirical content (Popper, 1966). Thus Popper tries to reduce the search for explanations to a search for high empirical content which is in turn reduced to a

¹⁷ Essentially the same point has been made by Rudner, who writes for example ". . . systematization seems to me as much a desideratum of science as is truth ..." (1961, p. 118).

search either for elimination of error or for progression towards the truth.

Both these attempts to reduce the aim of seeking explanations to more fundamental aims fail for the simple reason that high empirical content *cannot be equated* with high explanatory power.¹⁸ For we can always increase the empirical content of a theory by tacking on independently testable postulates, and this could quite clearly drastically *decrease* the simplicity, the explanatory power, of the theory in that the number of logically independent postulates would go up (a point made in chapter 1 above).

My suggestion, then, is that we should take the search for explanations—for explanatory systematization—as a *fundamental* aim of science, and not as a *derivative* aim; not as an aim to be reduced to some more fundamental aim by means of some such claim as that explanatory theories are more verifiable, more falsifiable, or better candidates for high verisimilitude than non-explanatory theories. Explanatoriness should be conceived of as an end in itself, and not as a means to some other end.

The situation before us, then, is this. The problem that Popper fails to solve is:

- (a) Given that the aim of science is to progress towards the truth (or, more modestly, to eliminate error), how can we show that the methods of science give us the best, the most rational hope of realizing this aim?

¹⁸ High empirical content is at best a necessary condition for high explanatory power or simplicity; it is certainly not a sufficient condition. Actually Popper has in effect acknowledged the need for a criterion of simplicity, or explanatory power, which goes beyond the notion of empirical content. For he asserts that we should require of a new theory that it should "proceed from some simple, new, and powerful unifying idea" (1963, p. 241), and goes on to call this a requirement of simplicity, which cannot, it seems, be formulated very clearly. Popper does not, however, go on to argue (as I do in the text) that the search for explanations or simplicity should be taken as a fundamental aim of science, and that this enables us to provide a rationale for both the rules and the aims of science.

This problem needs, I suggest, to be replaced by the following problem:

- (b) Given that the aim of science is to develop theories of both increasing explanatory power *and increasing* verisimilitude, how can we show that the methods of science give us the best, the most rational, hope of realizing this aim?

Problem (a) is insoluble for the following fundamental reason. Physics places a premium on theories of extreme simplicity, of high explanatory power. No reason can be given, however, for supposing that *simple* theories as opposed to *complex* theories are either nearer to the truth or more falsifiable. If the world is in fact extremely complex so that in the end no explanation can be given for phenomena, then the simpler our theories become, the further from the truth they will be. Again, a *complex* theory of high empirical content is strictly just as falsifiable as a *simple* theory of equal empirical content. Thus, given either the aim of developing theories of increasing falsifiability (i.e. the aim of increasingly efficient error elimination) or the aim of approaching the truth, no rationale can be provided for those methodological rules which place a premium on *simple* theories, on theories of high explanatory power.

This fatal difficulty does not, however, arise in connection with problem (b). Given that a fundamental aim of science is to search for explanations no difficulty arises in providing a rationale for those methodological rules which favor *explanatory* theories to non-explanatory theories.

This easy victory is of course won at a price. Problem (b) is solved at the expense of creating a new problem. For, of course, in order to exhibit science as rational enquiry (and thus adequately solve the demarcation problem) we need to show not only that the rules of science are rational, but also that the *aim* of science is a rational aim to adopt. We thus have the new problem:

- (c) How can we show that the aim of seeking theories of both increasing explanatory power (or simplicity) and increasing verisimilitude is a rational aim to adopt?

Before considering problem (c), however, I wish to discuss in a little more detail my solution to problem (b). We can, I suggest, argue as

follows. As long as we are pursuing the aim of seeking *explanatory truth*, it will be entirely rational to plan our strategy on the assumption that our search will meet with success. But our search can only meet with success if the world does ultimately have a simple structure, or, in other words, if a certain metaphysical thesis—let us call it the thesis of "structural simplicity"—is true. For if the world is ultimately incredibly complex, then as our theories increase in explanatory power, or simplicity, they will move further and further away from the truth. Therefore, as long as we seek *explanatory truth*, it will be entirely rational to base our strategy on the assumption that the thesis of structural simplicity is true.

It may be asked: But how can it possibly be rational to make such an incredibly risky assumption as the thesis of structural simplicity when we have no reason whatsoever to suppose this thesis to be true? That we have no reason to suppose the thesis to be true I grant. The crucial point however is this: As long as our aim is to seek explanatory truth it is only risky to make those assumptions which, if wrong, threaten to endanger the success of our enterprise. But in assuming the thesis of structural simplicity to be true we risk nothing since the truth of this thesis is just a condition for our enterprise to be successful. Making the assumption that the thesis of structural simplicity is true can in no way whatsoever endanger or block the success of our enterprise. If this assumption *is false*, then there is no way whatsoever in which we can realize our goal, whether we make the assumption or not. Thus making the assumption adds no additional risk to the success of our enterprise whatsoever.¹⁹ (It is only relative to the search for truth *per se* that the assumption becomes incredibly risky.)

¹⁹ One potential reason only exists for considering the possibility that the thesis of structural simplicity is false, granted that we are seeking explanatory truth. If there were the slightest hope of establishing that the thesis of structural simplicity is false, then it would be entirely rational to consider this possibility (so that we can call our whole enterprise off the moment we discover our goal is unobtainable). But there is no hope of establishing that the thesis of structural simplicity is false, any more than there is a hope of establishing that it is true. The best that we can do, by way of showing that the search for explanatory truth cannot succeed, is to pursue this goal as strenuously as possible, and after untold years of

Granted, then, that we are pursuing the aim of seeking *explanatory* truth we may entirely rationally, and wholly without risk, plan our strategy on the assumption that the thesis of structural simplicity is true, even though we have no reason whatsoever to suppose that this thesis is true. But if the thesis of structural simplicity *is* true, then it is only reasonable to suppose that the regularities we observe in the world—the low-level hypotheses that we have corroborated—are due ultimately to the structural simplicity of the universe. As physicists, in seeking more and more accurate delineations of the structural properties of the universe, our best policy must be to develop high-level, bold conjectures which explain as much of the observed regularities—the low-level corroborated hypotheses—as possible. When our high-level theory conflicts with a number of well-corroborated low-level hypotheses, and when our attempts at patching up our theory lead to considerable theoretical complexities, it is only rational to suppose that our original high-level theory is on the wrong lines. We must look for a new theory, involving fewer independent postulates, which can explain all that the old theory failed to explain. In particular, it will be rational to choose a new theory which successfully predicts a number of striking new phenomena: for such a theory holds the promise of delineating quite accurately some aspect of the ultimate structure of the universe.

This argument, if successful, only establishes the rationality of adopting roughly Popperian rules *granted that we are pursuing the aim of seeking explanatory truth*. But is this a rational aim to pursue in the first place? How, in other words, are we to solve problem (c)?

It may be thought that problem (c) is just as insoluble as problem (a). For if the aim specified in (c) is to be rational, do we not need to have some rational assurance that this aim will meet with success? But this aim of seeking explanatory truth can only meet with success if the thesis of structural simplicity is true. And no reason, no rationale, can be given for holding that the thesis of structural simplicity is true rather than

effort, fail in the attempt. Thus, granted that we are pursuing the goal of seeking explanatory truth, the only potential reason for considering the possibility that the thesis of structural simplicity is false collapses.

false. It looks, then, as if the aim of seeking explanatory truth cannot be a rational aim.

One standard way of attempting to solve the problem of induction, or the problem of providing a rationale for methodological rules, is to introduce a metaphysical thesis which, if true, would explain the success of our inductive inferences, the validity of our methodological rules (see Lakatos, 1968a for a discussion of this kind of approach). The obvious objection to such an approach is that it only succeeds if we can provide some reason for holding that the relevant metaphysical thesis is true. If no reason whatsoever can be provided for this, then the introduction of the metaphysical thesis does not help at all. And of course no reason can be given for maintaining that it is more rational to hold that such a metaphysical principle is true rather than false. Thus the familiar attempt to resolve the problem along these lines fails.

But does not the approach that I am advocating here fail for similar reasons? Does not the rationality of the aim of seeking explanatory truth require that it is more rational to hold that the metaphysical thesis of structural simplicity is true rather than false?

The answer is: no. *An aim can be rational even though we have no rational assurance whatsoever that the aim will meet with success.* Of course, if we have rational reasons for believing that an aim *cannot* meet with success, then it is irrational to pursue such an aim. But in the absence of such reasons, it may well be rational to pursue some goal even though we have no reasons in advance for maintaining we will meet with success. This view is enshrined in the adage that one cannot know what one can do until one tries. Perhaps the majority of new projects, initiated by mankind, have been initiated without advance rational assurance of success. Of course if a project meets with no success whatsoever after long effort, then gradually this in itself may be taken as a reason for holding that the project *cannot* succeed, and is irrational. (Perhaps alchemy might be viewed in this light.)

Thus for the aim of seeking *explanatory truth* to be rational it is not necessary that it is more rational to hold that the thesis of structural simplicity is true rather than false. All that is required is that it is not positively irrational to hold that the thesis of structural simplicity might *possibly* be true. It is only if it were more rational to hold that the thesis of structural simplicity *is false* rather than true that the aim of seeking explanatory truth would fail to be rational.

Thus the approach advocated here successfully overcomes the standard lethal objection to the familiar attempt to provide a rationale for methodological rules by introducing an appropriate metaphysical principle. Problem (c) is solvable even though problem (a) is not.

One final objection, to my claim that the aim of seeking explanatory truth is rational, needs to be considered. It might be objected that I have, in a sense, established too much in that I have established the rationality of any crazy line of enquiry whatsoever.

Consider for example the marvelous fairy tale of scientology. At one time we were all gods. But we grew bored. So we created the universe, and immersed ourselves in it, to keep boredom at bay. But after several reincarnations, we gradually forgot who we were and the powers that we possessed, until we ended up in our present pitiful state, as mere human beings. Therapy can, however, enable us to recover our memory and our powers.

Now a scientologist can, it seems, argue that the aim of developing successful therapeutic methods is just as rational as the physicist's aim of seeking explanatory truth. What entitles us to claim that the aims of physics are more rational than the aims of scientology?

Ultimately I think we can only say that it is only the enormous apparent success of physics which entitles us to make this judgment. (I say *apparent* success, for of course we do not know that the aim of seeking explanatory truth, via physical enquiry, has met with *real* success.) The moment scientology began to meet with a *comparable degree of apparent success* the scales would begin to tip the other way. If a scientologist, after therapy, was able, merely by the exercise of thought, to shift the planets in their course round the sun, or alter a law of nature or two, then physics might begin to look rather silly as an alternative rational search for explanatory truth.

Chapter Three

Science, Reason, Knowledge, and Wisdom: A Critique of Specialism

(First published in *Inquiry*, 23, 1980, pp. 19-81)

I

In this chapter I argue for a kind of intellectual inquiry which has, as its basic aim, to help all of us to resolve rationally the most important problems that we encounter in our lives, problems that arise as we seek to discover and achieve that which is of value in life. Rational problem-solving involves articulating our problems, proposing and criticizing possible solutions. It also involves

breaking problems up into subordinate problems, creating a tradition of specialized problem-solving – specialized scientific, academic inquiry, in other words. It is vital, however, that specialized academic problem-solving be subordinated to discussion of our more fundamental problems of living. At present specialized academic inquiry is dissociated from problems of living – the sin of specialism, which I criticize.

I proceed by discussing two rival views about the nature of intellectual inquiry. I call these two views *universalism*²⁰ and *specialism*. I shall argue that at present the whole institutional structure of scientific, academic inquiry, by and large, presupposes specialism. Of the two views under consideration it is, however, universalism, and not specialism, which provides us with a rational conception of intellectual inquiry. Failure to put universalism into practice has profoundly damaging consequences for science and scholarship, and indeed for life, for our whole modern world. Ideally

²⁰ When this chapter was first published, in 1980, I called universalism *fundamentalism*. Unfortunately, the latter term has been hijacked to mean religious dogmatism and primitivism. I have therefore jettisoned the term “fundamentalism” and adopted “universalism” instead – although I will, on occasions, continue to speak of such things as “critical fundamentalist inquiry” where I mean “universalist inquiry”. As the argument develops, it will emerge that the position I wish to defend ought really to be called *specio-universalism*, rather than just *universalism*. It is, in a sense, an admixture of *specialism* and *universalism*. Or rather, it ought really to be called *critical specio-universalism*, but that is too much of a mouthful to repeat throughout the chapter. I should add that, having selected “universalism” in the hope that I now have a term free of the unfortunate religious connotations of “fundamentalism”, I was somewhat dismayed to find that the Oxford Concise Dictionary defines “universalism” as the Christian view that everyone will be saved. This has the merit of being more humane than those views which consign many to eternal damnation. Nevertheless, in what follows, please ignore this Christian interpretation of the term. “Universalism”, here, means what I say it means.

intellectual inquiry ought to help us to tackle rationally those problems of living which we encounter in seeking to discover and achieve that which is of value in life. Intellectual inquiry ought, in other words, to devote reason to the enhancement of wisdom (wisdom being defined here as the capacity to discover and achieve that which is of value in life, for oneself and others – wisdom thus including knowledge and understanding, but much else besides). In fact, at present, scientific, academic inquiry gives priority to the achievement of knowledge only, rather than to the achievement of wisdom. It is essentially the general adoption of specialism which is responsible for the persistence of this highly undesirable state of affairs.

II

According to universalism, in the end the whole point of intellectual inquiry is to help us to improve our answers to four universal – or fundamental – questions, namely:

- (1) What kind of world is this?
- (2) How do we fit into the world and how did we come to be?
- (3) What is of most value in life and how is it to be achieved?
- (4) How can we help develop a better human world?²¹

²¹ These fundamental problems may of course be formulated a little differently from this without affecting the overall argument. I shall argue, in fact, that these problems need to be understood, at the most fundamental level, as personal and interpersonal, or social, problems which we encounter in our lives. The exact form in which problem (3), for example, arises for any individual will depend upon the circumstances in which the individual finds himself. 'How can I get enough to eat?', 'How can I find worthwhile, productive work to engage in?', 'How can I give and receive love?', 'How can my life be of value if I am to grow old and die?', 'What am I to do with my life?', 'How can I develop my present pursuits so that I achieve more successfully that which is of real value?' These can all be regarded as possible variants of problem (3).

In particular, according to universalism, it is a basic task of intellectual inquiry to help us to tackle these four fundamental problems in *a rational* fashion. Rational problem-solving is understood here to involve, at the very least, putting into practice the two heuristic rules:

(a) Articulate, and seek to improve the articulation of, the problem to be solved;

(b) Propose and critically assess possible solutions.²²

There is of course more to rational problem-solving than this.²³ But these two rules are understood by universalism to constitute the nub of rationality.

Thus, according to universalism, the central and fundamental task of intellectual inquiry is to improve the articulation of the above four problems, and to propose and critically assess possible solutions to them. All other intellectual activity is subservient to this.

A basic idea of universalism is that ideally it is we ourselves who answer the above four questions, as we live. The proper task of reason, of thought, of intellectual inquiry is to help us to arrive at answers that we really do wish to give to these questions, rather than to determine the answers for us. Intellectual inquiry is our servant, not our master. It is not in itself any kind of authority or oracle.

²² ' . . . the one method of all *rational discussion* . . . is that of stating one's problem clearly and of examining its various proposed solutions *critically*' (Popper, 1959, p. 16).

²³ We may regard a problem as having the form of an aim we seek to realize and some provisional idea for a route to the realization of our aim, which fails, however, to enable us to achieve the aim. As a result of representing problems in this fashion, we may well adopt the idea that rationality involves quite essentially seeking to improve our aims and methods as we act by imaginatively developing and critically scrutinizing possible and actual aims and methods. For an exposition of this somewhat more sophisticated 'aim oriented' conception of rationality – and its implications for intellectual inquiry – see chapters 4, 5, 7 and 8, and my (1976, esp. ch. 9, and 1984, chs. 3 and 4).

Two further extremely important, elementary heuristic rules of rational problem-solving are:

(c) Break up the basic problem to be solved into subordinate, specialized, easier-to-solve problems.

(d) Interconnect attempts to solve basic and specialized problems, so that the one may influence and be influenced by the other.²⁴

According to universalism, an immense amount of intellectual activity arises, quite properly, as a result of putting these two heuristic rules into practice. That is, in order to improve our answers to our four basic problems we create a vast network of sub-problems and preliminary problems – the specialized, technical problems of science and scholarship. A great deal of intellectual activity consists in seeking to solve these limited, technical problems of specialized scientific, academic disciplines. It is however of supreme importance – according to universalism – that we do not

²⁴ Rules (a) and (b) specify *universalism*, pure and simple. It is the addition of rules (c) and (d) that transforms *universalism* into what may be called *specio-universalism*: see note 1. *Specialism* results when rule (c) alone is implemented, rules (a), (b) and (d) being ignored, at least as far as fundamental problems are concerned.

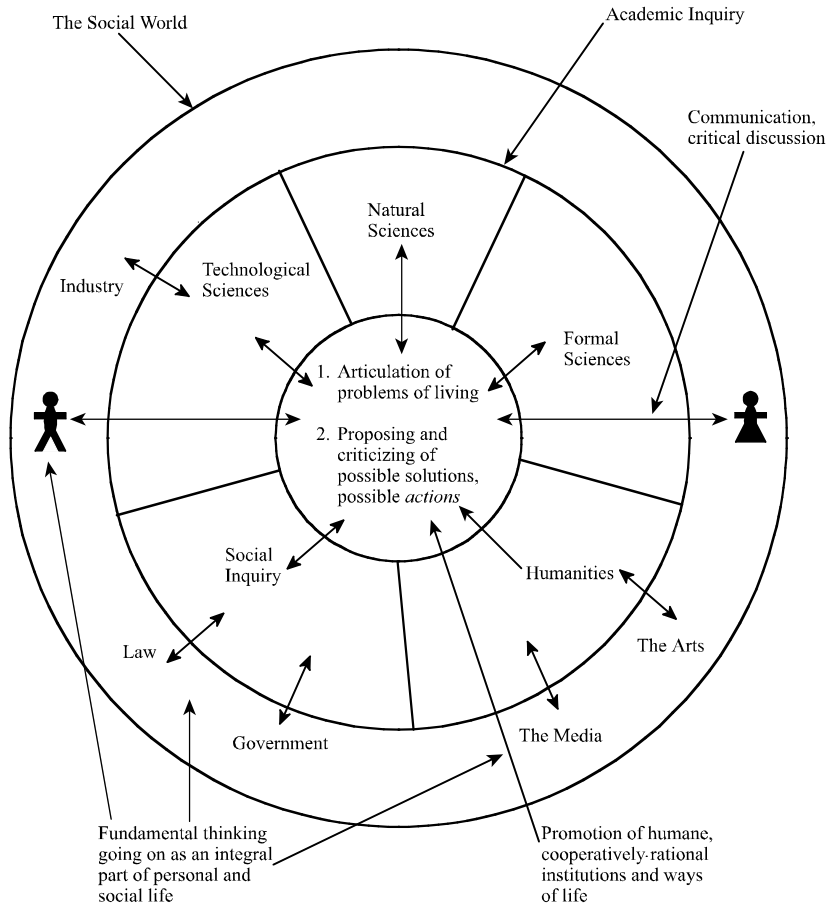


Diagram 1: Specio-Universalist Academic Inquiry Helping People Realize what is of Value in Life

lose our way within this network, this maze, of sub-problems. If intellectual inquiry is to be rational, it is essential that intellectual priority be given to the four fundamental problems, and to the tasks of proposing and critically assessing possible solutions to them. In order to tackle specialized problems in a rational fashion, in short, it is essential to tackle such problems as sub-problems of the four fundamental problems. Specialized scientists and scholars, in other

words, in order to be rational, must also be philosophers or generalists, concerned in their specialized work to help us solve our fundamental problems.

Diagram 1 gives an indication of the way in which some current specialized academic disciplines may be conceived, in critical fundamentalist terms, as being designed to help us solve the above four basic problems. As the diagram indicates, it is essential for the intellectual integrity and rationality of intellectual inquiry as a whole that there be a constant two-way flow of information between specialized problem-solving and fundamental problem-solving.

Two minor adjustments may be made to the doctrine of universalism as just outlined. In the first place it may be argued that philosophy ought not to be conceived as yet another specialized discipline concerned to solve its own special problems. Rather, philosophy needs to be conceived as that part of the whole intellectual enterprise which seeks to articulate fundamental problems, propose and criticize possible solutions to these problems. Philosophy, according to this conception, constantly gives rise to new specialized problems, and is itself profoundly influenced by our success and failure in seeking to solve specialized problems. It is in just this sense that almost all the great philosophers of the past have contributed to 'philosophy': Plato, Aristotle, Descartes, Bacon, Hobbes, Locke, Spinoza, Leibniz, Hume, Kant, Mill, Comte, Marx, Rousseau, Nietzsche, Mach, Russell – to name a few. It is vital, however, according to this viewpoint, that philosophy is not treated as a specialized, professional discipline, the exclusive preserve of the expert. The whole rationale of intellectual inquiry is to promote fundamental rational problem-solving as widely as possible, as an integral part of life. Rendering this the exclusive task of professional philosophers sabotages utterly the whole *raison d'etre* of intellectual inquiry. This non-specialized, critical fundamentalist conception of philosophy is perhaps above all to be found upheld by the thinkers of the Enlightenment – for whom critical philosophy was the basic instrument of human enlightenment.²⁵

²⁵ See P. Gay (1973, vol. 1, pp. 3-19, 127-203).

Strictly, of course, formulating universalism in terms of this Enlightenment conception of philosophy requires us to modify Diagram 1, in that 'philosophy' ceases altogether to be any kind of specialized academic discipline existing alongside other disciplines, and becomes instead identical to all thought about fundamental problems, ideally pursued rationally as an integral part of life. Formulating universalism in this way, however, is unfortunately liable to lead to misunderstandings. Academic philosophers will object to the disappearance of specialized philosophy, not realizing that there must always be an important place in academic inquiry for those concerned with fundamental problems and concerned to promote open, critical discussion of fundamental problems. Everyone else will object to the idea that philosophy should monopolize concern with fundamental problems, 'philosophy' being misunderstood here to mean 'academic philosophy' rather than being understood to be simply all our personal and public thinking about our fundamental problems. The essential tenet of universalism after all is that *all* inquiry, personal, social, and academic, ought to be organized along universalist lines. Whether or not thought about fundamental problems is called 'philosophy', and whether or not academic philosophy continues to exist as a specialized discipline, are matters of minor importance. In order to avoid misunderstandings concerning these minor matters, I leave Diagram 1 unmodified.

The second qualification that may be made to the above viewpoint amounts to this. It is quite wrong – it may be argued – to suggest that the enterprise of seeking to improve our answers to the above four fundamental questions is somehow exclusively the concern of intellectual inquiry. Literature, theatre, music, art, religion can all be interpreted as being concerned to illuminate our responses to these basic questions – especially the last two questions. Our whole culture can, in other words, be conceived of in universalist terms as being designed, ideally, to help us to discover and create that which is of most value in life. In engaging in our work, in social and political activity, we should, ideally – it may be argued – be seeking to develop improved answers in practice to the last two questions, in one way or another. Indeed, in

our whole way of life – our way of being on this earth – we give implicitly our actual answers to such questions, whether we are aware of this or not. And in so far as we seek to improve our lives, we seek to improve the actual answers that we give to these questions, in the fabric of our actions. Universalism, in short, needs to be conceived as a philosophy of life, a social philosophy, a philosophy of culture: universalism interpreted as a philosophy of intellectual inquiry is simply a fragment of all this.

Universalism, as just characterized, may seem at first sight to be a somewhat autocratic, doctrinaire position, in that it seems to determine for us what our problems are and how they should be conceived. For this reason, it may at first sight seem unacceptable. For do not our problems – even our ‘universal’ or fundamental problems – change, quite legitimately, from circumstance to circumstance, from person to person, from culture to culture? Can we really ever know for certain what our fundamental problems are, how they should be conceived?²⁶

It will I hope become clear, as the argument unfolds, that my basic purpose in this chapter is to depict – and argue for – a kind of intellectual inquiry specifically designed to offer us maximum help with discovering for ourselves, whoever we may be, what our own unique problems of living are, how we are to conceive of them, and how we are to set about resolving them. My claim is that intellectual inquiry, so designed, is universalism. It is intellectual inquiry so designed that it has the kind of intellectual-institutional structure depicted in the diagram above, according to which problems and their discussion are, as it were, hierarchically organized, with four vague, general, fundamental problems at the top, a maze of specific, restricted, precise, specialized problems at the bottom, and in between a continuous range of problems, more and less specific, inter-connecting the top and the bottom by means of the relationship: ‘Problem P_1 is more fundamental than problem P_2 ’ or, equivalently, ‘Problem P_2 is subordinate to problem P_1 ’. A few universal, fundamental problems are needed so that we do not get lost in the maze of restricted, specialized

²⁶ I am grateful to L. Briskman for provoking me into discussing this objection explicitly.

problem-solving. These fundamental problems must be formulated informally, imprecisely, without restricting specific presuppositions, just so that all people everywhere, in all societies, cultures, and circumstances, can in principle interpret their own more or less specific, basic problems as specific versions or interpretations of the four fundamental problems, as formulated above. Only this can ensure that no one is excluded *a priori* from entering into rational inquiry by their own specific circumstances, view of the world, philosophy of life. In addition, we need discussion of more precise, restricted problems so that we can make progress with solving our problems, as a result of putting into practice the third and fourth of the above four rules of rational problem-solving, (c) and (d).

Universalism needs to be implicit in the way in which our own personal thinking and problem-solving is organized, so that we may have the best opportunity to understand and learn from others, even from those who think very differently from ourselves – learning from others being essential for the development of our own capacity to recognize and solve our own problems.²⁷ Universalism

²⁷ I even put this forward as a psycho-neurological hypothesis: our wonderful unconscious problem-solving capacity, which we exhibit so effortlessly in life whenever we perceive, understand, speak and act, is due to the fact that a fundamentalist hierarchical structure is programmed, as it were, into the neurological structure of our brains. This has evolved as a result of natural selection (problem-solving ability – and above all the ability to solve relevant problems, procured by the fundamentalist hierarchical structure – having great survival value). Unfortunately, at present, nothing like so intelligent a structure is built into scientific, academic inquiry – or into much conscious thought – in that here, lamentably, specialism prevails. In particular we have failed to build the hierarchical structure of universalism into our civilization. Not surprisingly, this civilization, or world order, at present exhibits a terrifying failure to recognize and resolve its fundamental problems – problems most relevant to the achievement of what is of most value-even to the extent that its very survival is now in doubt.

needs to be built into education, into the intellectual-institutional structure of scientific, academic inquiry, and generally, into our whole social, political, economic, and cultural order, on a world-wide basis, so that learning, understanding and cooperation between people is given every opportunity to flourish.

There is nothing autocratic or doctrinaire in what I am advocating here, just because universalism amounts to a kind of intellectual inquiry, a way of thinking or problem-solving which, when put into practice, gives us our maximum chances of discovering for ourselves what our own unique problems are and how they are to be solved, enabling us, ideally, to exploit for this purpose the very best thinking or problem-solving that humanity has to offer. The autocratic and doctrinaire, the dogmatic, arise to the extent that *we fail* to put universalism into practice.

Critical fundamentalist intellectual inquiry can thus incorporate all possible conceptions of the world, all religious views, all philosophies of life, in all possible social and cultural milieux – all possible ways of conceiving of life's problems and how they should be tackled. There is just one proviso: all these diverse views and values, in being plugged into critical fundamentalist inquiry, as it were, must take note of the following basic points: many ways of conceiving of the world, life and its problems, exist and are possible; whoever we may be, our view as to what sort of world this is, and what is of most value in life, is guesswork; we have much to learn from others – especially by taking the achievements and failures, the views, values, and arguments of others seriously, by ourselves engaging, with others, in critical fundamentalist inquiry, as we live. Sincere attention to the lives, views and values of others is desirable – and ought to be held to be desirable – within all viewpoints and value-systems, since this is absolutely essential for mutual understanding in the world, mutual learning, mutual cooperation, peace, friendship, and love. Much of the real richness in life comes from the good things that go on between people; and for these good things to happen, sincere attention to the lives, views, and values of others – universalism built into the pattern of our lives, the structure of society – is essential.

Universalism takes into account the point, stressed especially by Popper in *The Open Society and its Enemies*,²⁸ that social, cultural pluralism or diversity is essential for the development of reason and science – the development of what Popper has called *critical rationalism*. I shall even argue, somewhat analogously to Popper, that rational inquiry can be understood as developing as a result of our departure from tribal life – from the human compactness and unity of tribal life. In sharp disagreement with Popper, however, I wish to argue that such things as mutual cooperation, mutual learning, understanding, and communication can only flourish within social and cultural diversity if some kind of common unity can be discovered within this diversity. We must be able to agree at some level about what sort of world this is, and what is desirable and of value. Engaging in cooperative intellectual inquiry – the very act of participating in rational discussion – presupposes that it is at least possible to discover or create, at some level, common purposes and assumptions, an agreed framework, an agreed outlook on life and the world. This agreement must, however, accommodate equably the existing differences. It is in order to do justice to this requirement of unity in diversity – essential for cooperative rational discussion and inquiry – that universalism postulates or stipulates the above kind of hierarchical ordering of problems and their attempted resolution. The hierarchical structure of critical fundamentalist inquiry is precisely what we need if we are to discover or create, as readily as possible, just, equitable agreement within disagreement, unity within diversity. Agreement can be sought at the fundamental level: disagreement and doubt can then be rationally explored at less fundamental levels, wherever it arises.

In his best epistemological, social, and political thought Popper is centrally concerned to attack authoritarianism, the dogmatic attitude. In *The Open Society and its Enemies* this concern takes the form of a mighty onslaught upon those major figures in the history of Western thought who, in Popper's view, have failed to come to terms with the strains of civilization – the strain of living in an open, pluralistic society – and, as a result, have given way to

²⁸ Popper 1966a).

romantic longings for the cohesion of the closed society, the tribal way of life. It is this longing, this potent false nostalgia for a golden past, which Popper argues has led even some of the greatest minds, with the best of intentions, to become the enemies of the open society, the enemies of democracy, reason and pluralism, and as a result, tragically, actually helping totalitarianism and fascism to grow, with all the consequent appalling human suffering of our history.

Popper's diagnosis is of fundamental importance. However, in the midst of his ferocious determination to establish once and for all the intellectual disreputability and appalling potential human destructiveness of views which value the tribal way of life, Popper neglects to consider the possibility that there is indeed much to value, potentially, in the cohesiveness of the tribal way of life which humanity – science, reason, and civilization – cannot do without.²⁹ It

²⁹ Some modern writers have done full justice to the great potential value of living and working in a small community or 'tribe': see, e.g., Turnbull (1976), Schumacher (1973), and "A Blueprint for Survival", (1972). Popper's failure to recognize this potential must be due partly to his being unacquainted with the anthropological evidence. He asserts that 'the main element' of the tribal 'magical attitude towards social custom' is 'the lack of distinction between the customary or conventional regularities of social life and the regularities found in "nature" 'this often being associated with 'the belief that both are enforced by a supernatural will'. Social customs are rigidly maintained, there being a superstitious fear of change, magical 'taboos rigidly regulating] and dominating] all aspects of life'. Significantly, Popper adds that 'comparatively infrequent changes have the character of religious conversions or convulsions, or of the introduction of new magical taboos': see Popper (1966a, vol. I, p. 172.) It is striking that Turnbull finds all these Popperian characteristics of tribal life dominating the life of agricultural Bantu tribes in central Africa. Turnbull describes just such a rigid, taboo-ridden, superstitious, compulsive, fearful, ritualistic way of life. Turnbull's really remarkable discovery, however, is that all this is entirely absent in the Pygmy hunting and gathering tribal way of life. The Pygmies' lives are imbued with a quite extraordinary spontaneity, grace, and trust, there being a complete

absence of superstition, compulsive ritual, or fearful observance of taboo. Turnbull argues, in my view entirely convincingly, that it is the development of agriculture which is responsible for this dramatic difference in the whole way of life. Hunting and gathering tribes can afford to live spontaneously, from day to day, trusting in the forest to provide food for tomorrow. Agricultural tribes, on the contrary, live in a state of constant battle with the environment and must perform persistent, long-term agricultural work before food and reward are eventually forthcoming. [M. Harris, in his (1978), comes to the conclusion, from a consideration of archaeological evidence, that early hunting and gathering tribes "enjoyed relatively high standards of comfort and security", having more leisure than later agricultural tribes.] Thus, it is not closeness to Nature, but the exact opposite, departure from day-to-day dependency on Nature, the development of agricultural technology, which creates rigidity, taboo, and ritual. In any case, the Pygmies decisively refute Popper's contention that tribal life is invariably rigid, ritualistic, and irrational. In many ways, in fact, our modern 'open' societies in the industrially advanced West are closer, at the institutional level, to the Bantu reliance on rigidly maintained ritual and taboo, than to the Pygmy reliance on spontaneous instinct and skill. And – of particular relevance to the theme of this essay – this is perhaps especially true of modern specialized scientific, academic research. Rigidly maintained taboo and ritual, broken only by 'comparatively infrequent changes' having 'the character of religious conversions or convulsions' – this corresponds almost exactly to specialist scientific research as described and documented by Kuhn (1970a). Even the vocabulary is the same. Kuhn describes scientific revolutions as infrequent episodes of crisis, inducing intense anxiety while they last, the process of acquiring the new paradigm constituting a kind of irrational religious conversion.

At present one perhaps needs the serene self-assurance and lucidity of an Einstein (acquired as a result of sustained, instinctive fundamentalist thought) to recapture the spontaneity and trust of the Pygmy way of life in the modern scientific world. It is clear that in Einstein's case scientific curiosity arose spontaneously from

is just this possibility that is affirmed here. I shall argue that our departure from the human compactness and unity of tribal life does indeed involve serious loss. Mere pluralism is not enough. It is essential that we develop a common unified view of the world and ourselves through cultural and social diversity if there is to be mutual learning, understanding, and cooperation through diversity – minimal requirements for reason and for civilization. Only universalism can do justice to these apparently conflicting requirements of unity and diversity. In our emergence from tribal life into the modern world a basic task confronting us is to create and develop unity within diversity: only by putting universalism into practice can we achieve this in a just, equitable, genuinely rational and humane way. Popper's ideal of the 'Open Society' needs to be replaced by the ideal advocated in this essay of the 'Fundamental Society'. It is precisely our failure to establish universalism on a world-wide basis that is responsible for so much suffering in modern times, and which indeed at present threatens us all. (I refer here to our present world-wide incapacity to cope with fundamental problems posed by such things as the population explosion, the continuing rapid depletion of vital, finite natural resources, widespread poverty and malnutrition in the third world, global warming, the proliferation of nuclear weapons which threatens to engulf us all in the nuclear holocaust. A critical fundamentalist world order is, almost

the heart in response to a feeling of 'rapturous amazement at the harmony of natural law'. [In a letter to Gertrud Warschauer in 1952 Einstein wrote: 'You have given me great joy with the little book about Faraday. This man loved mysterious Nature as a lover loves his distant beloved. In his day there did not yet exist the dull specialization that stares with self-conceit through hornrimmed glasses and destroys poetry. . . .': Dukas and Hoffmann (1979, p. 42)]. And Einstein found no difficulty in conceiving himself as a part of Nature. When asked to respond to the question 'If, on your death bed, you looked back on your life, by what facts would you determine whether it was a success or failure?', Einstein replied: 'Neither on my death bed nor before will I ask myself such a question. Nature is not an engineer or contractor, and I myself am a part of Nature' (Dukas and Hoffmann, 1979, p. 92).

by definition, a world order capable of recognizing its fundamental problems and, where possible, developing and putting into practice, in a cooperative fashion, just, humane, effective solutions.)

Popper's failure to recognize the vital need to create or develop a version of tribal unity within the diversity, complexity, and sheer immensity of the modern world, in order to preserve and develop reason, mutual cooperation, humanity and civilization, is intimately connected with his analogous failure to recognize the vital role that fundamental unifying assumptions play in science, and in academic inquiry in general. Scientific, academic inquiry has basic presuppositions about what sort of world this is and what is important or of value in social life, built into its whole intellectual-institutional structure, built into the priorities for research, built into its implicit methodology. According to universalism, these basic presuppositions need to be explicitly articulated and scrutinized – thus creating a tradition of discussion of presupposed solutions *to fundamental* problems – if scientific, academic inquiry is to be genuinely rational and rigorous, of maximum human value and use. Only by putting universalism into scientific, academic practice can we do justice to – and develop – the inherent rationality, the intellectually progressive character, and the human value, of the best of scientific, academic work and thought. As we shall see below, Popper fails to characterize adequately the rationality and progressive character of science – in that, for example, he fails to solve the problem of induction – just because he fails to do justice to the need for fundamental metaphysical and evaluative assumptions persisting through scientific revolutions, scientific diversity.

The Open Society and its Enemies fails to characterize a genuinely rational society: *The Logic of Scientific Discovery* fails to characterize a genuinely rational science: both failures are by-products of Popper's basic failure to articulate and advocate the hierarchical structure of universalism, so essential for genuinely rational, cooperative problem-solving and inquiry in life as well as thought.

Having argued that we need to recognize, quite generally, that our thinking goes on in the world, presupposing a view of the world and a view of what is of value in life, I am of course eager to

acknowledge that my advocacy of universalism in this essay is intimately bound up with a view of what sort of world this is and what is of value in life – a broad, general, fundamental answer to the fundamental questions (1) and (3) above. As to the material universe I hold a view not too dissimilar from the overall conception of the world implicit in much of modern science – a view of the world which does justice to the probable truth of Einstein's remark that 'all our science, measured against reality, is primitive and childlike – and yet it is [one of] the most precious thing[s] we have'.³⁰ I recognize, of course, the intellectual legitimacy of conceptions of the world – such as animistic and religious views – very different from that of modern science: critical fundamentalist inquiry acknowledges such rivals, and retreats to a more modest 'common sense' view of the world, designed to be indifferent between these rival, explanatorily fundamental views, so that there may be a common, agreed base in terms of which the merits of the rival explanatorily fundamental views may be discussed. (universalism recognizes, in other words, that, in certain contexts, and for certain purposes, the epistemologically fundamental may differ from what is presumed to be ontologically and explanatorily fundamental.)

As to that which is of value, I hold that all that is of value in existence has to do with life, and especially, for us, with human life. Enjoyment in living; curiosity and wonder; perceptive awareness, understanding and appreciation of significant and beautiful aspects of the world; kindness, laughter, honesty, friendship, love, intimacy, cooperative creative work, personal responsibility, happiness, fulfilment: these are the kind of things that are of value. For each one of us, this short life is our only opportunity to discover, experience, take part in, life of value; all too many people in the modern world – especially the third world – lack this opportunity. We need to do all we can to change things so that all people everywhere have the opportunity to realize what is of most value in life. Value in the world has much to do with the diversity of life, the unique particularity of each individual life. A uniform world would be a world denuded of value. It is of the essence of

³⁰ From Hoffmann (1973, p. vii).

value – it is essentially desirable and of value – that there be multiplicity and variety, amongst people, amongst ways of life, amongst societies and cultures. However, if this desirable variety is to flourish in this one crowded world, it is essential that we discover how to cooperate, to learn from and understand each other, in the midst of this variety. And there is a further point. As I have remarked above, much of what is of value in life comes directly from good things that go on between people – mutual understanding and appreciation, sharing, intimacy, cooperative creative work. Such interpersonal or social things, of value in themselves, only become possible in a world full of variety if there is cooperation, communication, learning, and understanding amidst variety. Variety is only enriching in so far as there is understanding and learning between people amidst variety. It is to help facilitate all this that I advocate universalism (or critical fundamentalism). Universalism is put forward as a conception of learning and problem-solving designed, above all, to help us resolve more adequately the third and fourth of the above fundamental problems.

Amongst other things, universalism amounts to a reply to social and cultural *relativism*. Like relativism, universalism acknowledges the existence and value of social, cultural, and intellectual diversity. Unlike relativism, universalism recognizes that we all live in a common world in which we all have a real value, and that we all need to learn from one another so that mutual understanding and cooperation may flourish – so that what is of value in all our lives, potentially and actually, may flourish. The existence of a multiplicity of cultures need not prevent us from recognizing our common humanity, our common value, since it is at least possible for this multiplicity to be interlaced with and unified by a common acceptance of universalism.

Adoption of universalism is especially important for societies and cultures in the third world. For in learning from the industrially advanced West – in acquiring the science, technology, and industry of the West – there is always the grave danger that the indigenous culture and social order will simply be

annihilated,³¹ as opposed to being helped to develop and flourish. A third-world society can only avoid this danger by articulating, at a fundamental level, basic presuppositions, values, and problems of the society, so that it becomes possible to discover how to develop these presuppositions and values, solve these problems, in the new social and cultural circumstances made possible by the importation of Western ideas and techniques. Only in this way can such a society employ these ideas and techniques discriminatingly, for its own best purposes, instead of becoming a hollow imitation of the Western way of life. In addition, of course, the industrially advanced West has a special responsibility, in its interactions with the third world, to construe its own social and cultural order in universalist (or critical fundamentalist) terms. Only cooperative universalism can enable a mutually desirable kind of learning to go on in both directions.

Analogous considerations arise in connection with education. The most profound, instinctive, and passionate fundamentalist thinkers are of course very young children – since all children must, as a practical necessity, arrive at working answers to the four fundamental questions in order to become human. If education is to develop, and not annihilate, instinctively fundamentalist childish thought, then education must itself be organized along universalist (or critical fundamentalist) lines.³² Only those teachers who learn from their pupils really educate.

³¹ The destructive impact of industrially more advanced ways of life on primitive or so-called primitive ways of life has been, and is at present, all too often, blatant and brutal. But it can also be subtle and unintended. For a perceptive account of this in connection with the importation of Western economic ideas and practices, see Schumacher (1973).

³² Einstein was always aware of the instinctively fundamentalist character of childish thinking – as well as of the childish origins of mature fundamentalist thought – associated, for him, essentially with curiosity provoked by a sense of wonder, together with scepticism concerning the received dogmas of the adult world. In explanation of his own fundamentalist thinking concerning the structure of the physical universe he once remarked that ordinarily

To sum up: critical fundamentalist inquiry does justice to the Socratic and Kantian idea that reason forms a basis for the unity of

only children take such problems seriously. He, however – a late developer – continued to pursue such elementary questions; and, as an adult, naturally, was better equipped to come up with improved answers. On another occasion he remarked: 'There exists a passion for comprehension, just as there exists a passion for music. That passion is rather common in children, but gets lost in most people later on. Without this passion, there would be neither mathematics nor natural science': Einstein (1973, p. 342). And in connection with his own education, in a well known passage, he remarks: 'In this field ... [of physics] I soon learned to scent out that which was able to lead to fundamentals and to turn aside from everything else, from the multitude of things which clutter up the mind and divert it from the essential. The hitch in this was, of course, the fact that one had to cram all this stuff into one's mind for the examinations, whether one liked it or not. This coercion had such a deterring effect [upon me] that, after I had passed the final examination, I found the consideration of any scientific problems distasteful to me for an entire year. In justice I must add, moreover, that in Switzerland we had to suffer far less under such coercion, which smothers every truly scientific impulse, than is the case in many another locality. There were altogether only two examinations; aside from these, one could just about do as one pleased. This was especially the case if one had a friend, as did I, who attended the lectures regularly and who worked over their content conscientiously. This gave one freedom in the choice of pursuits until a few months before the examination, a freedom which I enjoyed to a great extent having gladly taken into the bargain the bad conscience connected with it as by far the lesser evil. It is, in fact, nothing short of a miracle that the modern methods of instruction have not yet entirely strangled the holy curiosity of inquiry; for this delicate little plant, aside from stimulation, stands mainly in need of freedom; without this it goes to wrack and ruin without fail. It is a very grave mistake to think that the enjoyment of seeing and searching can be promoted by means of coercion and a sense of duty': Einstein, 'Autobiographical Notes', in Schilpp (1969, p. 17).

mankind, in such a way as to encourage the flourishing of desirable kinds of diversity within this unity; it might be called 'the tribal discussion of humanity'. Universalism cannot, of course, of itself vanquish tyranny, exploitation, manipulation, war, terrorism, crime. Universalism does, however, hold out the hope that if it is actively promoted in our personal, social, intellectual, economic, and political lives wherever possible, then the spirit and practice of mutual cooperation between people may gradually grow, thus enabling us gradually to dismantle those social and cultural arrangements which tend to breed misunderstanding and mistrust, manipulation, and exploitation, the use and abuse of power, the dreadful spiral of threat, counter-threat, and violence.

So much for my preliminary exposition of universalism. I turn now to a consideration of the rival doctrine *of specialism*.

For most scientists, scholars and educationalists today, specialism is a much more familiar doctrine than universalism: my exposition of specialism can therefore be much briefer. Specialism, unlike universalism, is almost exclusively a view of professional, expert, scientific, academic inquiry – even though this view, being embodied in so much present-day scientific, academic practice, has far-reaching consequences for all our personal, social lives. In complete contrast to universalism, specialism insists that only the specialized, technical problems of the various academic disciplines deserve serious intellectual attention. In order to be capable of serious scientific or scholarly treatment, in other words, a problem must satisfy certain conditions. It must be capable of being given an agreed, precise formulation. The problem must have an objective character, in that experts agree as to how the problem is to be formulated. The nature of the problem must not depend on such subjective, personal, or idiosyncratic matters as mood, feelings, personal desires, attitudes, or convictions. There must exist agreed procedures for tackling the problem. Above all, there must be general agreement as to what counts as the solution. It must be possible for the problem to receive a definitive solution. Academically respectable problems must, in short, have many of the characteristics of puzzles – chess or crossword puzzles for example – as emphasized by Kuhn in connection with what he

has called 'normal' science.³³ Such problems arise quite essentially within the context of specialized disciplines, where there are agreed methods, results, assumptions, procedures. It is precisely by excluding all that is vague, ambiguous, controversial, metaphysical, or philosophical that such academically respectable problems can be formulated or created. In order to be in a position to understand, solve, and assess proposed solutions to such problems one needs to be an expert, with specialized knowledge of the relevant discipline, its methods and results. It is not in general necessary to have broad intellectual or cultural sympathies and understanding. By and large, ignorance of social, political, religious, moral, and philosophical issues lying beyond the scope of his discipline does not in any way hamper or disqualify the expert in his professional work. A 'mere' expert or specialist can be as well equipped as anyone to make outstanding contributions to his discipline.

Experts can be in a position to pronounce authoritatively and definitively on matters that fall within the field of their specialized knowledge. In addition, *only* experts can be in a position to make such authoritative pronouncements: the rest of us cannot legitimately challenge or criticize expert judgments unless we too have specialized knowledge. Scientists and scholars are thus fully justified in ignoring criticism of their work and results by 'outsiders', by those without expertise. The price that the expert pays, however, in being able to make unassailable, authoritative judgments is that he must confine himself, *qua* expert, to delivering judgments that lie within the limited sphere of his professional competence – that small part of his discipline about which he does have expert knowledge. He must not in his capacity as expert make pronouncements about broad political, moral, religious, and philosophical issues – the immensely complex human, social problems of real life – which, in their very nature, cannot be amenable to specialized, academic treatment.

Specialism may seem to represent an intolerably narrow-minded, dogmatic, scholastic conception of intellectual inquiry. All that is adventurous, imaginative, speculative, free-ranging, and creative

³³ See Kuhn (1970a, ch. IV).

may seem to be excluded from science and scholarship. Those who defend specialism, however, usually do so in terms of the following kind of argument. It is precisely by eschewing consideration of imaginative, speculative, imponderable issues, and instead concentrating attention on much more limited, specialized 'puzzles', capable of definitive solutions, that science and scholarship have made such giant steps forward in recent times. In the end, sustained attention paid to limited, technical problems pays dividends, and may even result in a definite solution to some 'profound' philosophical problem. The problem of how the human race has come into existence has been discussed fruitlessly for centuries. Not until the work of Darwin was any real contribution made towards solving this 'philosophical' problem. The crucial point about Darwin's contribution, however – so the argument goes – is that it arose out of painstaking attention to highly detailed, limited, specialized problems within zoology and botany.³⁴

It is, I hope, obvious from the above that according to specialism the four basic problems of universalism lie wholly outside the field of reputable science and scholarship. Inevitably these four problems are such that there can be no general agreement as to how they ought to be formulated, or what methods ought to be adopted in seeking to solve them. It is most improbable – perhaps even undesirable – that there should ever be general agreement as to what is to count as a correct, acceptable solution to any of these problems. And it is extremely unlikely that any of them will receive a definitive solution. The four basic problems of universalism satisfy none of the requirements which specialism demands of academically reputable problems. Thus, according to specialism, discussion of these four problems has no place at all within scientific, academic inquiry. Academic inquiry may perhaps produce work that has some bearing on the answers we give to the four basic questions, as in the case of Darwin's work. This comes about, however, as a result of aiming at solutions of exclusively specialist, technical problems. The four

³⁴ Kuhn, e.g., argues that the instigation of the specialized, autonomous puzzle-solving of the specialist is essential for scientific progress. See Kuhn (1970a, pp. 21, 24, 37, and 64-65).

basic problems of universalism only have a place within academic inquiry at one remove, as it were, within anthropology, sociology, or the history of ideas. A historian of ideas, for example, may quite legitimately discuss the writings of those who have speculated about such problems. Such a historian will however be concerned to solve specialized problems within his field, concerning the evolution of ideas. He will not concern himself with the fundamental problems as such – not if he is to continue to function as an intellectually reputable academic.

Extreme versions of specialism – such as logical positivism – condemn the four basic critical fundamentalist problems as metaphysical and evaluative, and therefore strictly meaningless. Less extreme versions of specialism merely place them outside the domain of intellectually respectable scientific, academic inquiry.

III

Universalism and specialism uphold diametrically opposed intellectual standards.

According to universalism, it is absolutely essential for the rationality, intellectual rigour, and integrity of intellectual inquiry as a whole that sustained attention be given to the four basic problems. Indeed, this attention needs to be given intellectual priority over all else. All other intellectual activity needs to be subservient to the central and fundamental activity of imaginatively proposing and critically assessing possible answers to the four basic problems. Only in this case can even the most elementary of requirements for rational problem-solving be realized.

According to specialism, on the other hand, rationality, intellectual rigour, and integrity, actually demand that the four 'basic' problems of universalism be placed outside the domain of reputable intellectual inquiry. Mature science, authentic scholarship, genuine intellectual progress only really get underway when inconclusive philosophical debate about fundamental issues has been put firmly aside.³⁵

³⁵ As we shall see below, there is a further vital point of difference. Universalism asserts that inquiry can only be really intellectually

One important aspect of this difference in intellectual standards is that universalism and specialism uphold different conceptions of intellectual progress.

According to universalism, intellectual progress is to be conceived in terms of the success that intellectual inquiry has in enabling us to improve our answers to the four fundamental problems, and to improve our capacity to tackle these problems in a rational fashion. One might say that universalism, ultimately, conceives of intellectual progress in personal and social terms – in that what is at issue is the answers that people in fact give to fundamental questions in their lives. Our assessment of intellectual progress will of course depend to some extent on the kind of tentative, broad answers that we give to these questions. Intellectual progress itself is no doubt something absolute and definite; our assessment of intellectual progress, however, is bound to be somewhat tentative, it being possible for there to be a number of different legitimate assessments.

According to specialism, on the other hand, intellectual progress is to be conceived in terms of the success that intellectual inquiry meets with in solving specialized, technical, scientific/academic problems. Progress – or the lack of it – is thus something definite, uncontroversial, something about which there can be general agreement. This is especially true for science. According to specialism, all scientific problems are essentially problems we encounter in seeking to predict more and more phenomena more and more accurately. Thus scientific progress is to be assessed simply in terms of the success we meet with in developing laws and theories which predict more and more phenomena more and more accurately.³⁶

rigorous if it is recognized that inquiry (thought, problem-solving), at the most fundamental level, goes on in life as an integral part of our personal and social lives, actively helping us to discover and achieve what is of most value in life, potentially and actually, as we live.

³⁶ A further clarification, to be elaborated below. Universalism conceives of intellectual progress, fundamentally, in personal and social terms, in terms of progress in our achievement of what is of value in life, in terms of the progress in our personal and social

IV

Actual scientific, academic inquiry, as it exists at present, and has existed during the last hundred years or so, amounts to an uneasy admixture of universalism and specialism. In many ways, however, specialism predominates.

It must of course be acknowledged that some aspects of scientific, academic inquiry do exemplify critical fundamentalist standards. For example, there can be no doubt that science, technology and scholarship have made great progress when viewed from an universalist (or critical fundamentalist) perspective. The special and general theories of relativity and quantum theory have changed profoundly our conception of the physical universe. The theory of evolution, and subsequent developments since Darwin's day, have done much to improve our understanding of how we fit into the world and have come to be. Our whole conception of the cosmos has been utterly transformed during this period. Technological research has done much, potentially and actually, to provide us with the means to create a better human world. Research in history, archaeology, anthropology – and more questionably, research in other social sciences and humanities – has deepened our understanding of ourselves, our past, our potentialities.

In addition to this there have been many noteworthy 'universalist' or 'critical fundamentalist' thinkers who have consciously sought to help solve one or other of the four fundamental problems. Almost at random one might mention: Einstein, Freud, Schrödinger, Eddington, Russell, Whitehead, Poincaré, Jung, Erich Fromm, Margaret Mead, Karl Popper, Carl Sagan, E. Schumacher, I. Illich, T. Szasz, F. A. Hayek, A. Koestler, T. Roszak, H. Marcuse, R. May, R. Higgins,³⁷ and of course many others of varying repute.

thinking actively associated with and guiding our endeavours to achieve what is of value, on a personal and world-wide basis.

³⁷ A remark about the first and last of these 'fundamentalists'. Einstein once said: 'I want to know how God created this world. I'm not interested in this-or-that phenomenon, the spectrum of this-or-that element. I want to know His thoughts, the rest are details':

In many ways, however, the influence of universalism on actual scientific, academic practice is submerged beneath the massive influence of specialism on all but a minute proportion of scientific, academic work. Most scientists and scholars are specialists, concerned only to solve specialist problems not consciously conceived of as sub-problems of the four fundamental problems. Almost all scientific, academic publications are concerned with the resolution of specialist problems. Education is shaped primarily by specialist assumptions and standards, especially towards the upper end of the educational ladder, culminating as it does in the extreme specialism of the Ph.D. thesis. Academic appointments, academic honours, academic success, are all judged in terms of specialist standards – apart from quite exceptional cases.

Perhaps most crucially of all, the overall organization, the institutional structure, of scientific, academic inquiry exemplifies specialism rather than universalism. Universities are split up into relatively autonomous faculties: for example, faculties of physical sciences, biological sciences, technology, medicine, humanities or arts. Each faculty is subdivided into a number of relatively autonomous departments corresponding roughly to distinct academic disciplines. On the intellectual level, however, the subdivisions proceed further: each discipline is subdivided into a number of sub-disciplines: a specialist whose field of expertise lies within such a sub-discipline may not even be able to communicate properly – let alone share problems – with colleagues working within the same

see Salaman (1979, p. 22). In *The Seventh Enemy: The Human Factor in the Global Crisis*, R. Higgins (1978) outlines with devastating clarity and force six basic threats to the future of civilization – six fundamental world-wide problems which we must somehow resolve on a world-wide basis if mankind is to survive. His 'seventh enemy' is our human incapacity to acknowledge and respond to these fundamental problems, on both individual and social, political or institutional, levels. Thus, on a world-wide-basis, life on earth is at present almost lunatically irrational in the most elementary fashion (since it fails to put into practice the two most elementary rules of rational problem-solving).

discipline. Such an expert will communicate almost exclusively with his fellow specialists scattered throughout the world – thus participating in what has been called an 'invisible college'.³⁸

The striking point to note about all this is that nowhere is any provision made whatsoever for sustained, explicit, influential discussion of fundamental problems. This does not exist at the level of individual universities; nor does it exist at the level of published intellectual discussion, at the level of 'invisible colleges'.³⁹ Scientific, academic inquiry is, in other words, organized overwhelmingly in accordance with the intellectual standards of specialism.

All this has dire intellectual consequences – especially, of course, if viewed from the perspective of universalism. The remorseless concern to solve exclusively specialist problems for their own sake, the proliferation of specialized disciplines (disciplines within disciplines, the autonomy of each jealously guarded), the accumulation of specialized results and vocabulary, increasingly specialized education (specialist indoctrination), the absence of informed, critical, non-technical discussion of fundamental issues – all these factors combine to make it overwhelmingly difficult for

³⁸ See Price (1961). It must be emphasized that this modern meaning of the phrase, introduced by Price, is a typical specialist perversion of the original fundamentalist meaning intended, e.g., by Boyle in the seventeenth century when he writes: The "Invisible College" consists of persons that endeavour to put narrow-mindedness out of countenance by the practice of so extensive a charity that it reaches unto every thing called man, and nothing less than an universal good-will can content it. And indeed they are so apprehensive of the want of good employment that they take the whole body of mankind for their care. But. . . there is not enough of them': quoted in Werskey (1978, p. 13).

³⁹ It is noteworthy, e.g., that Higgins (1978) is obliged to break all conventional academic boundaries in order to articulate our basic global problems. It is also noteworthy that these problems discussed by Higgins and others do not receive sustained, influential discussion as an integral part of the orthodox scientific, academic enterprise.

anyone to discover, understand, and use the fundamentalist implications of specialized results. Intellectual inquiry becomes increasingly fragmented and incoherent, increasingly unusable from the standpoint of helping us to improve our answers to the four fundamental questions.

That over-specialization can have undesirable consequences has, it is true, been rather widely recognized. This scarcely amounts, however, to a recognition of the inadequacy of specialism. For if we look at what has been done in an attempt to compensate for fragmentation brought about by over-specialization, we find that new interdisciplinary subjects have been created, subjects such as biophysics, biochemistry, mathematical logic, industrial sociology. This typically specialist way of attempting to solve the problem actually, in many ways, serves only to make it worse. In seeking to facilitate communication between disciplines, additional buffer disciplines are created which only have the effect of further obstructing interdisciplinary communication. Thus even those who seek to combat some of the bad consequences of specialism can only adopt specialist methods in seeking to do so – so powerful a hold does specialism exercise over the academic mind – the end result being in consequence the exact opposite of what was originally intended. What cannot be done, of course, is what is needed most: the development of a tradition of influential, informal discussion of fundamental problems, feeding into, and being fed by, diverse specialist discussion. This obvious solution cannot be adopted for the simple reason that it involves violating specialist intellectual standards!

A further powerful indication of the increasing predominance of specialism over universalism is provided by the way in which academic philosophy has developed in recent times. Increasingly, academic philosophers have been concerned to develop philosophy as an academically respectable specialized discipline, with its own particular problems and methods, existing alongside other academic disciplines. For the vast majority of academic philosophers, progress in philosophy is to be achieved by pinpointing and solving technical problems mainly conceived as problems of 'conceptual confusion'

requiring 'conceptual analysis'.⁴⁰ Universalism, of course, becomes quite impossible if 'philosophy' is pursued in this specialized way.

⁴⁰ I refer here, of course, to the dominant schools of philosophy in Britain and the USA since the war, ordinary-language philosophy, conceptual analysis, logical empiricism, and descriptive metaphysics, as practised by, e.g., Ryle, Austin, Ayer, Anscombe, Warnock, Hare, Kenny, Strawson, Carnap, Hempel, Quine, Davidson, and many others. Recently there have been indications of some improvement in this tradition. Thus Mary Midgley's recent book *Beast and Man: The Roots of Human Nature* (1979) can be interpreted as making a valuable contribution to the fundamental problem: How do we fit into the world and how have we come to be? Unfortunately, Midgley takes for granted the conceptual analyst's conception of philosophy, and a form of specialism for intellectual inquiry as a whole.

In contrast to this tradition there are, e.g., Marxist philosophy and existentialism. Unfortunately, Marxist philosophers are more or less committed to interpreting philosophy as conscious or unconscious ideology – Marxist philosophy, in particular, thus being the detailed, specialized development and application of Marxist social theory and ideology – an attempt, by intellectual means, to help humanity realize Marxist social and political objectives, prejudged to be desirable. As a result, Marxist philosophers fail to practise universalism, which involves, amongst other things, the conscious articulation and criticism of ideologies, social theories, social and political programmes interpreted as possible solutions to our fundamental problems. (Unlike Higgins, Marxists do not begin with problems, but rather with a basic, presupposed solution. As a result, radical critics of the Marxist solution cannot be valued by Marxists as colleagues concerned to help solve essentially the same problems, but must inevitably be judged to be ideological and political opponents. The cooperative, rational development and appraisal of alternative possible solutions to our problems – including Marxist solutions – thus becomes impossible.)

Existentialism, on the other hand, can be interpreted as insisting that our most fundamental problems are problems we encounter in

For universalism requires the existence of the Enlightenment conception of philosophy – philosophy conceived as the open, non-professional, unspecialized discussion of fundamental problems, influencing and being influenced by specialized problem-solving in all other scientific, academic disciplines. In seeking to develop academically respectable, professional, specialized philosophy, academic philosophers have sabotaged almost all possibility of developing intellectual inquiry in critical fundamentalist directions.

Consider the following specialist account of the way in which intellectual inquiry has developed over the centuries.

'Intellectual inquiry begins with myth, religion, and philosophy. Originally, philosophy (or perhaps theology or metaphysics) is the

our lives – problems of living. If so, then the version of universalism advocated in section VII amounts to a kind of thoroughgoing, radical, rationalist existentialism. Three features of this version of universalism, in marked contrast to some features of traditional existentialism, must nevertheless be stressed. (1) All scientific, academic problems are to be interpreted as rationally subordinate to our problems of living, as we live. (This stands in marked contrast to some strands of existentialism, to be found in Kierkegaard and Dostoevsky for example, which are romantic, anti-scientific, and anti-rationalist in character.) (2) Our fundamental problems of living are interpersonal, or social, in character, as well as personal or individualistic. (Again, this is in marked contrast to the exclusively individualistic emphasis to be found in the existentialist writings of Kierkegaard and Dostoevsky.) (3) Rational problem-solving quite generally – and thus rational personal, social problem-solving in particular – involves quite essentially retaining a record of past successes and failures. Rational problem-solving is essentially accumulative and progressive in character. (This is in marked contrast to Sartre's hysterical repudiation of the past. Enhancing our freedom, our capacity to discover and achieve what is of value in life, requires that we learn from the successes and failures of the past, our personal and social history. The Sartrian repudiation of the past, in order to achieve freedom, in fact *enslaves* us to compulsive acting out of impulse.)

queen of the sciences, other intellectual disciplines having only a highly subservient, specialized role to play within philosophy. This state of affairs exists in the thought of ancient Greece, in the thought of Mediaeval Europe, and, to some extent, in the thought of seventeenth-century Europe during the so-called scientific revolution. For Kepler, Galileo, Bacon, Descartes, Newton, Spinoza, and Leibniz, philosophy and theology represented the primary, central disciplines – so much so that science was known as “natural” or “experimental” philosophy. Gradually, however, successive disciplines emerged out of philosophy, dissociating themselves from the parent discipline of philosophy, intellectual success and progress being essentially bound up with this long process of dissociation. Over the centuries philosophy has given birth to the autonomous disciplines of mathematics, astronomy, physics, logic, biology, history, political science, sociology, psychology, cosmology, linguistics (the last three or four only having become autonomous in the twentieth century). As a result of having bred these autonomous disciplines, philosophy itself has been left in a highly impoverished state. The nature and status of philosophy, in other words, have changed dramatically. Instead of being the queen of the sciences, overarching all other sciences, philosophy has been transformed into a highly specialized, technical, somewhat meagre enterprise, concerned not with improving our knowledge and understanding of the world – for that is the business of the empirical sciences – but rather with clarifying concepts and solving conceptual problems. In line with the general trend, academic philosophy seeks to transform itself into a specialized discipline, dissociated from "philosophy" in the original sense of Plato, Spinoza, Leibniz, Diderot, Voltaire, Hume or Kant.⁴¹

It must be admitted, I think, that this specialist account of intellectual history does considerable justice to the way intellectual inquiry has *in fact* developed over the centuries. Furthermore, this account is today in practice widely upheld throughout the scientific, academic world as providing us with an adequate account of how

⁴¹ Something like this account is presupposed, or propounded, by: G. Ryle (1967); Ayer (1969, ch. 1, 'On Making Philosophy Intelligible', pp. 1-18); Whiteley (1955, pp. 5-6).

intellectual inquiry *ought* to develop. Scientists and scholars have had something like this account in mind in pursuing and developing diverse disciplines. Above all, most contemporary academic philosophers take for granted the conception of modern philosophy that emerges from this account.⁴² All of which provides a strong indication of the extent to which specialism has come to be built into the institutional framework of contemporary scientific, academic inquiry.

Universalism, of course, provides us with a quite different picture of how intellectual inquiry ought to develop. If intellectual inquiry begins with myth, religion, philosophy, metaphysics, this is because intellectual inquiry begins quite properly with a concern with the above four fundamental questions. Intellectual progress requires, of course, the development of specialized disciplines concerned to solve diverse subordinate and preliminary problems. It is of crucial importance, however, according to universalism, that this development occurs in such a way that we can, all the more readily, tackle the four fundamental questions in a rational fashion. The development of *autonomous* disciplines – the essential feature of the specialist account – violates the most elementary rules of rational problem-solving.

None of the above, however, captures that feature of present-day scientific, academic inquiry which constitutes the most blatant and harmful institutional embodiment of specialism. This feature concerns, not so much the internal intellectual-institutional structure

⁴² An amusing indication of this is the way in which philosophers tend to acknowledge, apologetically or critically, that philosophy still concerns itself with the problems discussed by, e.g., Plato, whereas other disciplines successfully solve initial problems and move on to new problems, thus making progress. The failure of philosophy to progress in this way is only problematic if philosophy is conceived in specialist terms. From the standpoint of the fundamentalist or Enlightenment conception of philosophy, it is of course precisely the basic task of philosophy to keep alive, throughout the whole of intellectual inquiry, and throughout our culture and social life, a sustained concern with our four fundamental problems.

of scientific, academic inquiry, but rather the way in which scientific, academic inquiry is related to society, life, and the problem-solving that goes on in all our personal and social lives. According to the version of universalism that I wish to defend, the basic task of professional scientific, academic inquiry is to help all of us to recognize and resolve rationally those problems we need to resolve in order to discover and achieve that which is most desirable and of value in life. The basic task of critical fundamentalist academic inquiry is to help us to put universalism into practice in our personal, social lives, and to help us to develop a social order, a world, in which cooperative rational resolving of our most important personal and social life-problems may receive every encouragement. For this goal to be realized, there must be a constant two-way flow of ideas and arguments between discussion of fundamental problems in society, as a part of life, and discussion of fundamental problems within professional scientific, academic inquiry. An intimate, two-way, rational relationship needs to exist between society and science, life and scholarship.

At present this vital rational socio-cultural relationship scarcely exists anywhere. This is largely due to the prevalence of specialism which *prohibits* the above rational social relationship. Specialism demands precisely that scientific, academic inquiry, in order to be intellectually rigorous, must be such that the intellectual domain of scientific, academic inquiry is decisively *dissociated* from the discussion of problems that goes on in society, as a part of life. Scientists and academics, upholding specialist intellectual standards, have done their utmost to develop and preserve this dissociation – in order, from their own standpoint, to preserve rigorous intellectual standards. As a result, the scientific, academic community has betrayed its most profound intellectual purpose (as seen from the perspective of universalism): to help us develop more rational, wiser ways of living, a more rational, wiser world. The result of this betrayal, not surprisingly, is that the production of specialist knowledge flourishes, while wisdom in life, world-wide wisdom, falters.⁴³

⁴³ Since this was first written, in 1979, some academics have become concerned that members of the public should contribute to

Of the two views under consideration, it is universalism, and not specialism, which provides us with a rational, intellectually rigorous conception of intellectual inquiry.

In assessing the relative merits of the competing doctrines of universalism and specialism, it is vital to recognize that universalism fully acknowledges the immense value of – indeed the absolute necessity for – specialized scientific, academic work and thought. It is often only by putting into practice the two basic rules of rational problem-solving (c) and (d), formulated above in section II, that it is possible to make any headway with improving our solutions to our fundamental problems. Specialized problem-solving, specialized scientific, academic work is absolutely essential, according to universalism, for rational problem-solving in general.⁴⁴ The decisive additional point insisted on by universalism is that it is absolutely essential to put into practice rules (a) and (b) too. There must be a sustained rational discussion of our common, fundamental problems both within the scientific, academic community and within society, intimately inter-connected with specialized scientific, academic problem-solving if intellectual inquiry is to serve our best interests in a genuinely rigorous, rational fashion. It is legitimate, even desirable, that many individual scientists and scholars be absorbed by the pursuit of highly restricted, specialized topics and problems. What is

discussion concerning science policy options. Thus, in the UK, the Royal Society produced a report on the future of nanotechnology, the result of a collaboration of scientists and non-scientists. The Economic and Social Research Council has funded a research programme, Science in Society, which has explored issues having to do with public engagement with science. There is a rather general recognition that communication between science and the public should go in both directions. It is not good enough for scientists merely to inform the public about science. These developments do not, however, amount to academia as a whole engaging with the social world in a two-way discussion about how our problems of living are to be tackled in increasingly cooperatively rational ways.

⁴⁴ It is this feature of universalism which led me to suggest the view should really be called “specio-universalism: see note 1.

vital is that the overall intellectual-institutional structure of scientific, academic inquiry and of society itself accord with the kind of hierarchical structure required by universalism – sustained, explicit attention being given to fundamental problems. Failure to put into practice – to institutionalize – this vital critical fundamentalist perspective must inevitably lead to the fragmentation and trivialization of intellectual inquiry, and to a general incapacity to tackle cooperatively and effectively mankind's fundamental problems. The institutionalizing of specialism, however, obliges us to neglect the critical fundamentalist perspective. As a result we cease to tackle rationally just those problems it is most important for us to tackle rationally. While diverse sub-problems may be brilliantly tackled, our most general and important problems fall into neglect.

The motivation for insisting that it is of the essence of rationality to articulate our basic problems, and to propose and criticize possible solutions, is really very simple. If we do this, we give ourselves the best chance of seeking to solve those subordinate problems which are relevant to our main objectives. If we do not do this, the chances are that we will become engaged in seeking to solve sub-problems which are entirely misconceived or wholly irrelevant from the standpoint of achieving our basic objectives. Putting specialism into practice, in other words, is almost bound to lead to a mass of problem-solving activity which is misconceived or irrelevant from the standpoint of what matters most in life – a fair comment, I suggest, on a great deal of scientific, academic inquiry as pursued at present.

What if no serious doubts really arose as to how we should answer the fundamental questions: What kind of world is this? How do we fit in? How have we come to be? What is of most value in life? How can we help develop a better human world? In that case universalism would be somewhat redundant. Serious doubts presumably would only arise in connection with much more specific, particular issues. But this is not our situation. The above questions are all profoundly problematic, even if many people appear not to recognize the fact. Our greatest uncertainties simply do arise in connection with our most general and important problems. This being the case, it is essential that we give intellectual priority to the critical discussion of these problems, granted that we seek to develop a genuinely rational kind of intellectual inquiry.

Specialism is thus to be rejected, on the grounds that it provides us with a conception of intellectual inquiry that is both irrational and humanly undesirable, these two features indeed being intimately connected. Instead of prompting us to attend to what is most important and problematic, specialism does precisely the opposite!

The harmfulness of specialism does not lie in its tendency to encourage specialized puzzle-solving. Universalism, too, insists on the vital importance of such puzzle-solving. Nor need the harmfulness of specialism lie primarily in any tendency actively to suppress inquiry into fundamental problems. An upholder of specialism may simply see thought about fundamental problems as yet another specialized intellectual enterprise – grotesquely bankrupt intellectually, it is true, but scarcely deserving to be suppressed for all that. No, the real harmfulness of specialism arises from the fact that it appears to justify the pursuit of specialized problem-solving divorced from the consideration of fundamental assumptions and problems. Worse, specialism holds that intellectual integrity and respectability actually demand that fundamental assumptions – vague, conjectural, controversial – be excluded from specialized inquiry. As a result, the adoption of specialism leads to the development of specialized inquiries – within a multitude of diverse disciplines – all of which become immune to elementary, outside, fundamental criticism.

This feature of specialism is responsible for such widespread intellectual corruption in present-day scientific, academic inquiry, that it deserves further comment. The key point that needs to be recognized is that it must always be irrational and undesirable to pursue specialized problems isolated from all consideration of fundamental problems. This is because the whole paraphernalia of specialized problem-solving, as described above, actually requires us to give answers to fundamental problems. Choice of problems, formulation of problems, methods of attack, criteria for acceptable solutions, criteria for progress – all these essential features of specialized problem-solving implicitly presuppose more or less broad answers to the four basic questions – answers all too likely to be more or less false or unacceptable and standing in need of improvement. If specialized puzzle-solving cuts itself off from all critical consideration of fundamental issues (as specialism requires), then such puzzle-solving becomes irrational in the straightforward

and basic sense that implicit, influential, and controversial assumptions are made which are permanently protected from critical assessment. Only by openly acknowledging the basic metaphysical and evaluative presuppositions implicit in specialized puzzle-solving can such puzzle-solving become genuinely rational.

It is above all the enormous success of *science* – conceived of in traditional empiricist terms – which has seemed to provide the most powerful case for specialism, and for the central assumption that specialized problem-solving needs to be dissociated from fundamental assumptions and problems.

According to universalism, a basic task of science is to help us to improve our answers to the question: What kind of world is this? Thus, according to universalism, a genuinely rational science, putting into practice the two most elementary rules of rational problem-solving, gives intellectual priority to the task of proposing and criticizing answers to this question. Proposing and criticizing rival comprehensive metaphysical views about the nature of the universe, the nature of reality, constitutes, in other words, a central intellectual activity of a genuinely rational science.

Metaphysical assumptions at this level will influence drastically more restricted, specialized scientific problem-solving – the kind of methods adopted, the kind of theories developed and tested. Thus if we believe ourselves to be in some kind of animistic universe – or in an Aristotelian universe – we will adopt different methods and develop different theories from those which we will adopt and develop if we hold, in Galileo's words, that 'the book of Nature is written in the language of mathematics'. The success of modern science, according to this standpoint, is due in large measure to the fortunate choice of a comprehensive metaphysical conception of Nature – shared by Kepler, Galileo, and their successors – which sets the stage for a characteristic kind of specialized problem-solving. According to this critical fundamentalist standpoint, then, science needs to be understood in terms of an interplay between fundamental and specialized problem-solving, fundamental ideas and methods evolving with evolving specialized knowledge, this, in part, explaining the explosive growth of scientific knowledge. As our scientific knowledge improves, our knowledge about how to improve our knowledge – our *methods* – improves as well. All this

illustrates the four rules of rational problem-solving formulated above.⁴⁵

Just this critical fundamentalist conception of science – exemplifying elementary rules of rational problem-solving – is, however, rejected absolutely by almost all contemporary scientists and philosophers of science. For, according to traditional empiricist conceptions of science – almost universally taken for granted within the scientific community – it is the essential, defining characteristic of science that, in science, theories are selected impartially with respect to empirical success, independently of their compatibility or incompatibility with comprehensive metaphysical assumptions about the nature of the world. Many, of course, acknowledge that simplicity considerations play an important role in the assessment of scientific theories in addition to empirical considerations (for example Mach, Duhem and Poincaré); the decisive point, however, is that biased preference for simple theories in science is not interpreted as committing science to the metaphysical, and possibly false, assumption that the universe itself is simple. According to this traditional empiricist standpoint, science is successful precisely because theories are selected impartially with respect to empirical considerations isolated from all *a priori* metaphysical assumptions about the nature of the world. This was one of Bacon's main points. (Descartes disagreed: but with the downfall of Cartesian science, and the success of Newtonian science, generally and wrongly held to incorporate Baconian inductivism, Cartesian universalism was rejected by the scientific community.) The diverse philosophies of science of inductivism (Bacon and Mill), conventionalism (Duhem and Poincaré) and logical empiricism (Carnap, Hempel, and Nagel) all take for granted that in science theories are selected with respect to empirical success alone, unbiased by metaphysical assumptions about the nature of the universe as a whole. Even those thinkers who

⁴⁵ For a more detailed and sophisticated advocacy of this critical fundamentalist conception of science, see ch. 2, and my (1974; 1976a; 1977a; 1979; 1984; 1998; 2004; and 2007, especially chapter 14). For a critical assessment, see Kneller (1978, pp. 36-38, 80-37, 90-91); Muller (2008) and my reply (2009); McHenry (2009); Vicente (2010), Pandit (2010), and my reply (2010b).

acknowledge the importance of a *a priori* metaphysical ideas (Descartes, Spinoza, Leibniz, Kant) miss the essential point of the critical fundamentalist conception of science outlined above. For instead of emphasizing that our fundamental metaphysical ideas about the nature of the universe are *conjectures*, more or less bound to be false, and therefore needing constant critical scrutiny and development within science, these thinkers, on the contrary, seek to show, in one way or another, that fundamental metaphysical assumptions or principles can be conclusively established by reason, by argument. In effect empiricists and so-called 'rationalists' agree on one main point: metaphysical principles, unverifiable by experience, have a legitimate place in science only if they can be conclusively established by reason. Rationalists defend the existence of such principles: empiricists, correctly, reject this possibility. Both parties miss the essential point: metaphysical principles play a decisive role in science; these principles are, however, conjectures, more or less bound to be false: hence, if science is to be rational it is essential that these principles be articulated, criticized, and developed as an integral part of the scientific enterprise. Even Russell, it should be noted, misses this point. Russell recognizes that scientific method implicitly makes substantial metaphysical presuppositions about the world: he fails, however, to draw the critical fundamentalist conclusion from this, namely that a genuinely rational science seeks to improve its metaphysical presuppositions, and its methods, as it progresses.⁴⁶

The point is decisively rejected even by Popper. Popper has many critical fundamentalist arguments and remarks to his credit. His book *The Open Society and its Enemies* tackles an issue central to universalism. Popper emphasizes that metaphysical ideas have often

⁴⁶ See Russell (1948). Contrast Russell's uncritical or inflexible 'postulational' approach with the critical, flexible postulationism of aim-oriented empiricism, which stresses that science, in order to be rational, must continuously articulate, develop and criticize metaphysical blueprints for science as an integral pan of scientific inquiry, and in the light of ostensible scientific progress, thus enabling us to improve our aims and methods as our scientific knowledge and understanding of the world improves.

played a highly fruitful role in science.⁴⁷ He has emphasized the importance of 'metaphysical research programmes' for science, some science, in his view, even *amounting* to metaphysical research programmes (for example, in his view, the theory of natural selection).⁴⁸ He has argued that metaphysical ideas can be assessed rationally, as more or less adequate, tentative solutions to problems.⁴⁹ He has stressed that intellectual inquiry needs to be organized, not in terms of subject-matter and disciplines, but rather in terms of problems, and attempts to solve problems.⁵⁰ He has emphasized that science at its best is cosmology – the attempt, in effect, to answer the question: What kind of world is this?⁵¹ He has argued for philosophy conceived as a part of our attempt to improve our knowledge and understanding of the world, and against the view that philosophy is merely specialized 'puzzle-solving', or conceptual analysis.⁵² Finally, he has explicitly condemned specialism. Thus, commenting on the attitude of mind of the normal scientist, as described by Kuhn, Popper remarks:

I admit that this kind of attitude exists: and it exists not only among engineers, but among people trained as scientists. I can only say that I see a great danger in it and in the possibility of its becoming normal (just as I see a great danger in the increase of specialization, which also is an undeniable historical fact): a danger to science and, indeed, to our civilization.⁵³

Elsewhere he remarks:

If the many, the specialists, gain the day, it will be the end of science as we know it-of great science. It will be a spiritual

⁴⁷ Popper (1959, pp. 19, 38, 277-8).

⁴⁸ Popper (1976a, pp. 148-51, and sections 33 and 37). See also Popper (1983, section 23; 1982a, sections 20-28).

⁴⁹ Popper (1963, pp. 193-200).

⁵⁰ Popper (1963, pp. 66-67).

⁵¹ Popper (1959, p. 15; 1963, p. 136).

⁵² Popper (1963, pp. 67-96 and 136).

⁵³ Popper (1970, p. 53).

catastrophe comparable in its consequences to nuclear armament.⁵⁴

Nevertheless, the central tenet of Popper's thought in effect lends strength to a mainstay of specialism: namely, traditional empiricism. Much of Popper's later writings elaborate and apply the main thesis of his first book *The Logic of Scientific Discovery*. There Popper seeks to solve a problem central to traditional empiricism, namely how to demarcate science from metaphysics. Popper's solution, of course, is that a theory, in order to be scientific, must be experimentally falsifiable. A discipline, in order to be scientific, must assess theories solely with respect to empirical considerations, priority being given to those theories which have best survived severe testing and are most amenable to being severely tested. In other words, Popper, along with Bacon, Mill, Duhem, Hempel, and others, is centrally concerned to drive a sharp and decisive wedge between the assessment of specialized, partial solutions to scientific problems (laws and theories) and the assessment of solutions to the fundamental problem of science, namely metaphysical answers to the question: What kind of world is this? In *Conjectures and Refutations* Popper makes the matter altogether explicit when he defends 'the principle of empiricism which asserts that in science, only observation and experiment may decide upon the *acceptance* or *rejection* of scientific statements, including laws and theories'.⁵⁵ Dramatically and decisively, Popper rejects the basic tenet of the critical fundamentalist conception of science, as outlined above.

However, as I have argued at greater length elsewhere,⁵⁶ this 'standard empiricist' viewpoint is unacceptable. The insolubility of the problem of induction as formulated, for example, by Popper, shows clearly that scientific laws and theories – solutions to specialized scientific problems – cannot be assessed solely with respect to empirical success, in an entirely impartial fashion. If we honestly attempted to select theories in this way, we would always be overwhelmed by a vast number of complex, empirically successful theories, and we would fail to select the theories we do actually

⁵⁴ Popper (1976b, p. 296).

⁵⁵ Popper (1963, p. 54).

⁵⁶ See note 26

select in science. In practice, in science selection of theories is permanently biased in the direction of simplicity and unity, even to the extent of over-ruling mere empirical success. This means that in scientific practice, whether we recognize it or not, we presuppose that the universe has some kind of underlying structure (or at least that it behaves as if it had such an underlying structure, to a high degree of approximation). In other words, science is only possible in so far as a more or less specific, tentative answer is given to the question: What kind of world is this? Much of the success of modern science depends upon the aptness of this answer – so we may well judge. The answer is built into the whole methodology of science. In order to pursue science in a genuinely rational fashion, in a fashion which gives us the best hope of making real progress in improving our knowledge and understanding, we need to propose and criticize modified versions of our answer to the question 'What kind of world is this?' as an integral part of science. We need to do this in an attempt further to improve the methods, and the success, of science. We need in short to put universalism into practice. Any attempt, like Popper's, to characterize science in terms of fixed methods which select theories solely with respect to empirical success must fail to solve the problem of induction – simply because science, so characterized, violates the two most basic rules of rational problem-solving:⁵⁷ in addition, the

⁵⁷ This important point can be established quite simply as follows. Science is centrally concerned to solve the problem: What kind of world is this? If science is to tackle this problem rationally, priority needs to be given to proposing and critically assessing possible solutions – thus developing a tradition of rational cosmology like that represented, e.g., by Popper in his 'Back to the Presocratics' (1963, ch. 5). This leads, however, to the development of a number of rival imprecise possible solutions – rival cosmologies – with no indication as to how we are to make these vague ideas precise and choose between them. In order to proceed, we need to put into practice the third and fourth rules of rational problem-solving: each vague solution needs to generate preliminary, subordinate, specialized problem-solving. If one such approach begins to achieve apparent spectacular specialized success, then this entitles us to take this general approach especially seriously. Thus the

vital capacity of science to develop improved methods with improving knowledge, such an essential feature of scientific progress, must inevitably be missed out.⁵⁸ Instead of holding speculation about

spectacular specialized successes of Kepler and Galileo entitle us to take especially seriously their common vague cosmological presupposition: 'the book of Nature is written in the language of [simple] mathematics.' If science is to proceed rationally, however, it is essential that there continues to be an interplay between our best ideas as to how the over-all problem is to be solved, and our best solutions to subordinate problems. In particular, our assessment of possible solutions to subordinate problems – testable scientific laws and theories – must not be dissociated from our assessment of untestable, metaphysical ideas as to how the over-all problem is to be solved. Popper, however, violates this elementary, general requirement for rationality in insisting that assessment of scientific laws and theories *is* dissociated from assessment of metaphysical ideas. Furthermore, it is precisely this irrational insistence which creates, for Popper, the insoluble problem of induction. The impossibility of assessing scientific laws and theories solely with respect to empirical success is a special case of the general irrationality of attempting to assess possible solutions to subordinate problems independently of vague ideas about how to solve the over-all problem. The problem of induction, in short, is a product of specialism, the insolubility of the problem, as traditionally conceived, an indication of the irrationality of specialism as far as science is concerned.

⁵⁸ Ironically enough, Popper does come close to acknowledging the Russellian point that the methods of science make implicit metaphysical presuppositions about the nature of the world (see Popper, 1959. pp. 252-4), despite explicit disavowals elsewhere (see Popper, 1963, p. 54). He fails, however, to emphasize that critical rationalism requires that we explicitly articulate these metaphysical presuppositions, so that they may be criticized, and thus, we may hope, improved, as an integral part of science, so that the methods of science maybe improved with our improving knowledge. Just this way of doing science was instigated by Einstein in developing the special and general theories of relativity.

the ultimate nature of the universe to be metaphysical, philosophical, and thus of questionable scientific status, if not downright unscientific or even meaningless, we need, rather, actively to pursue such speculation, imaginatively and critically, as an integral part of science itself. We need to put into practice the kind of critical fundamentalist way of doing science so brilliantly initiated and exploited by Einstein, in developing the special and general theories of relativity.⁵⁹

That specialized scientific problem-solving requires some kind of answer to be given to the question 'What kind of world is this?'

The invariance and symmetry principles of modern physics – which can be interpreted as either methodological or metaphysical principles – are a development of Einstein's profound innovation. However, modern physics, and modern science quite generally, fail to put into practice, explicitly and fully, Einstein's way of doing science, in that they fail to articulate and criticize actual and possible aims and methods – or philosophies of science – as an integral part of science itself. The institutional reorganization that this requires – namely philosophy of science pursued as an integral part of science itself – has not been carried out. This is of course in part due to the fact that the scientific community accepts Popper's falsificationist demarcation criterion for dividing off science from non-science. Views about what ought to be the aims and methods of science – philosophies of science – not being themselves testable theories in any straightforward sense, have no place in science itself according to traditional, and Popperian, empiricism. Thus scientific integrity at present demands that discussion of aims and methods be excluded from science, instead of demanding that this discussion constitutes an integral part of science (as required by aim-oriented empiricism or universalism). At present, by and large, science departments and departments of history and philosophy of science do not speak to each other. To this extent Popper, rather than Einstein, is institutionalized. This institutionalization of Popperian methodology prevents us from developing a genuinely rational, fundamentalist science.

⁵⁹ For Einstein's advocacy of universalism see, e.g., Einstein (1973, part V). See also ch. 6 below.

has been vividly and dramatically demonstrated by Kuhn in his book *The Structure of Scientific Revolutions* (1970a). Kuhn establishes convincingly that the 'puzzle solving' of normal science depends upon the acceptance of a paradigm – in effect a *Weltanschauung*, a view of the world, for a given scientific discipline. One might well suppose that Kuhn, having realized this crucial point – this decisive objection to specialism – would go on to defend universalism, and the need for sustained development and criticism of 'paradigms' as an integral part of science. Kuhn, of course, does exactly the opposite. Discussion of fundamental issues has, for Kuhn, no place within a 'mature' science.⁶⁰ Furthermore, for Kuhn, changes of paradigm – scientific revolutions – inevitably involve a breakdown of rationality. Instead of emphasizing that rational assessment of paradigms is essential for the rationality of the whole of science – as universalism does – Kuhn, on the contrary, declares that choice of paradigm in general lies beyond the scope of reason.⁶¹ Kuhn, in short, is quite unable to conceive of non-specialist standards of rationality. *The Structure of Scientific Revolutions* brilliantly reveals the glaring defect of specialism, and yet, perversely, is itself a defence of specialism, of specialist intellectual standards. This provides yet another illustration of the powerful hold that specialism has over the academic mind – especially when one takes into account the great success of Kuhn's book in academic circles.

The profound irrationality of science as depicted by Kuhn in his book can perhaps be brought out by considering the following comparison. Our problem, let us suppose, is to wend our way through an obstacle-strewn path, from *A* to *B* (from ignorance to knowledge). Kuhn's advice is to proceed as follows. Standing at *A*, arrive at a general idea as to how to get to *B* (a paradigm); then, with head down – one might almost say with eyes shut – set off, sticking rigidly to this general idea. Even if you bump into a wall, fall into a ditch, or get tangled in brambles (anomalies), nevertheless adhere rigidly to your route (normal science).

⁶⁰ See note 15.

⁶¹ See Kuhn (1970a, ch. 12).

However, if you seem to have got into permanent difficulties (crisis), you may open your eyes, look around, and hit upon a new route (revolution), which, however, you must stick to as rigidly as before (new phase of normal science).

This blind blundering about may eventually bring you to your goal *B*. It is hardly, however, the most intelligent, the most rational way to proceed.

A rather more sensible procedure is to keep one's eyes open, and continuously adjust one's route (paradigm) in the light of what one sees and learns on one's way from *A* to *B*. In order to pursue science intelligently and rationally, in other words, we need to reconsider, explicitly and persistently, our most fundamental paradigmatic ideas as an integral part of science. Instead of adhering blindly and dogmatically to some paradigm until our difficulties have become overwhelming and we are forced to reconsider, we need rather to attempt to improve our paradigm even before insoluble empirical problems overwhelm us, taking into account important *a priori* considerations such as simplicity, coherence, unity, intelligibility, comprehensiveness. This was the way Einstein developed the special and general theories of relativity; Einstein was much too intelligent, and much too interested in discovering the 'thoughts of God', to follow Kuhn's advice.⁶²

⁶² It should be noted that the basic objection to Kuhn's prescription for science applies with almost equal force to Lakatos's prescription as outlined in his (1970). Lakatos's problem is to reconcile the dogmatism of Kuhn's normal science, on the one hand, with the anti-dogmatic, critical falsificationism of Popper, on the other hand, taking into account especially Feyerabend's important point that in order to test a given theory severely we need to possess, and even develop, alternative theories: see Feyerabend (1965). Lakatos's solution is to prescribe for science simultaneous competing fragments of Kuhnian normal science – competing research programmes – thus doing justice simultaneously to Kuhnian dogmatism and Feyerabendian pluralism. Lakatos makes it abundantly clear, however, that ultimately only relative empirical success ought to decide the fate of research programmes within science. There is thus no essential

To sum up this part of the discussion, specialized scientific problem-solving cannot proceed unless some kind of answer is given to the question 'What kind of world is this?'. This answer is almost bound to be more or less wrong, standing in need of improvement. Hence it needs explicit, sustained, critical discussion.⁶³

role, within Lakatos's conception of science, for sustained critical development of our best metaphysical answer to the problem. What kind of world is this?, so that the hard cores of research programmes could be assessed in part in terms of this answer. Lakatos advocates a kind of competitive specialism. In terms of our obstacle-course analogy, Lakatos sees science as a number of competing individuals, with different routes in mind, stumbling blindly from A to B.

⁶³ For a powerful criticism of the idea that the social sciences should be value-neutral, see Easlea (1973, pp. 167-78). Essentially the same point is made by Schumacher (1973), when he argues that economic thinking must reflect or presuppose some philosophy of life, some view as to what is of value in life. For the point that explicit articulation and criticism of value assumptions implicit in the aims of research is actually essential for the whole of science if it is to be objective and rational, see my (1977b, 1976a, chs, 5 and 7, and 1984). Values are even implicit, it should be noted, in the aims of a science as apparently remote from ordinary life as pure theoretical physics. The question: What kind of world is this?, may be interpreted in such a fashion that merely developing theories, like quantum theory, which predict more and more phenomena more and more accurately constitutes satisfactory progress towards answering the question. Einstein asked for much more from theoretical physics: he sought to capture, in a 'wildly speculative way' the 'thoughts of God'. He did not know that the universe has a coherent, unified structure: rather, the mere possibility of discovering such a structure seemed to him to be of such supreme value that to abandon the search for it seemed to be a profound betrayal of the noblest aspirations of theoretical physics. Thus Einstein's judgment that quantum theory is unsatisfactory, in that it abandons micro-realism, was in part based on a value judgment.

Specialized scientific problem-solving dissociated from such critical fundamentalist discussion is irrational, as our glance at Popper's and Kuhn's work has shown.

Analogous considerations arise in connection with all other specialized academic disciplines, and in connection with the other three fundamental problems. Inevitably, in pursuing specialized lines of research, in history, for example, in literary criticism, anthropology, sociology, psychology, medical research or engineering, we presuppose some kind of rough and ready answer to one or other – or to all – of the four fundamental questions, this answer influencing our choice of problems, criteria for successful solutions, and so on. Since such implicit and influential answers are all too likely to be more or less inadequate, it is essential, for rationality, that these answers be explicitly articulated and critically assessed, as an integral part of specialized problem-solving.⁶⁴

In recent years a number of writers – so-called 'externalist' historians of science and sociologists of knowledge – have argued in effect that specialized scientific, academic problem-solving is substantially influenced by the social and cultural circumstances in which it proceeds. Material conditions, religious, political, moral, and social ideals, human interests and values of one kind or another, all influence intellectual inquiry.⁶⁵ (This may be understood as a generalization of the Kuhnian point that specialized scientific problem-solving is influenced by paradigmatic assumptions.) Specialism insists that such non-rational influences must be kept to a

(For an endorsement of Einstein's judgment on this point see my 1976b; 1982; 1988; 1994; 1998, ch. 7) and chapter 13 below.

⁶⁴ A number of writers – e.g. Koyré (1973), Burt (1932), and Buchdahl (1969) – have advocated a view which might be called 'metaphysical presuppositionism', according to which the natural sciences do make substantial metaphysical presuppositions about the world. These writers fail, however, to emphasize the crucial point that scientific rationality demands sustained, explicit, critical development of such metaphysical presuppositions as an integral part of science itself – the essential tenet of universalism.

⁶⁵ Merton (1970); Mannheim (1952); Hagstrom (1974); Mathias (1972); Teich and Young (1973); Mulkay (1979).

minimum, and must be excluded altogether when results are being assessed, if intellectual inquiry is to retain its rationality and objectivity. Universalism, on the other hand, insists that such influences must be openly acknowledged and critically scrutinized if intellectual inquiry is to be rational and objective. If our task is to discover what is of value in life, and to help develop a better human world, then of course our thinking must not be dissociated from our personal and social lives, from our material circumstances, our political, moral, and religious ideals, our desires and values. A basic task of intellectual inquiry must be to promote more rational problem-solving in life – thus gradually helping us to develop a more rational human world: intellectual inquiry must not merely seek to shield itself from the corrupting influences of an irrational society, as specialism would have it.

Most contemporary externalist historians of science and sociologists of knowledge would probably agree that the specialist programme of excluding social and cultural influences from intellectual inquiry cannot succeed, and is even perhaps incoherent. One might suppose that as a result of recognizing the general untenability of specialism, these writers would advocate and practise universalism. In fact one finds nothing of the kind. Perversely, like Kuhn, these writers continue to accept and practise specialism – contributing to the highly specialized disciplines of history of science and sociology of knowledge. The main implication of this work is to undermine specialism: but if those who do this work do not themselves see this implication, how can anyone else be expected to see it? Once again we see the extraordinarily powerful hold that specialism has over the contemporary academic mind.⁶⁶

⁶⁶ On the one hand there are those who pursue sociology of science and 'externalist' history of science merely in order to add to specialist knowledge within sociology and history. These writers tend to decry the significance of epistemology and the study of scientific method. (A notable recent example of this is to be found in Bloor, 1976.) From the standpoint of the fundamentalist viewpoint defended in Sect. VII, this approach entirely misses the point. For, according to the view advocated below, the basic task of the social sciences is to help us develop more rational

institutions and ways of life, a more rational world. A central task of the social sciences, in other words, is to propose and critically assess possible institutional and social changes designed to help people all the better to discover and achieve what is of value in life – that is, to help people solve rationally the problems of living which they encounter in seeking to achieve that which is of value in life. The social sciences, on this view, ought thus fundamentally to be institutional or social epistemology or methodology. What is being attempted in this paper in connection with one institution – the scientific, academic enterprise – should be attempted quite generally in connection with institutions associated with politics, the law, the media, commerce, industry, and international relations. Far from the sociology of science taking over from the philosophy of science, sociology on the contrary – and the social sciences quite generally – need to become the philosophy and methodology of institutional, social pursuits and enterprises. Granted that our concern is to develop better solutions to problems (3) and (4), a central task of the social sciences and humanities ought to be to help us develop fundamentalist, or aim-oriented rationalistic, institutions quite generally – including aim-oriented rationalistic academic institutions. See my (1976a, chs. 8 and 9).

On the other hand, however, there are those Marxist-inclined writers who wish to commit science to socialist or Marxist objectives and who seek to 'radicalize' science. (See, e.g., Easlea, 1973; Rose and Rose, 1976.) These writers see social and cultural reality in terms of competing class interests – the dominant class ensuring that even culture and science serve its own class interests, this situation being maintained, in part, by means of the institutionalized lie that science is an objective, value-neutral search for truth, serving no special class interests. There is clearly some truth in this allegation. The moment we view scientific and technological research on a worldwide basis, it becomes clear that very little such research is devoted to serving the interests of the millions upon millions of desperately poor people in the third world. In so far as such research does serve social interests, it is the interests of those who live in industrially advanced, relatively wealthy countries which are served – even to the point of

increasing the misery of the underprivileged, as in the case, perhaps, of the tin miners of Bolivia. The fundamental defect of this Marxist conception of intellectual inquiry, however, is that it commits intellectual inquiry to socialist or Marxist social theory and objectives, and thus prevents intellectual inquiry from itself scrutinizing these social, political and evaluative presuppositions, even to the point of improving on them.

We might view the matter as follows. (1) Standard empiricists, like Hempel and Popper, reject the existence of permanent metaphysical presuppositions inherent in science. (Even Kuhn and Lakatos only allow for temporary metaphysical presuppositions to be assessed ultimately in terms of the empirical success of the specialist research they support; thus Kuhn and Lakatos ultimately also advocate standard empiricism.) This is dishonest, as the insolubility of the problem of induction indicates. (2) Metaphysical presuppositionists, like Russell, Koyré, Burtt, and Buchdahl, do acknowledge the existence of long-term, comprehensive metaphysical presuppositions implicit in science. This is more honest. These writers fail, however, to emphasize the crucial importance of articulating and critically developing such presuppositions as an integral part of science. In addition these writers fail to acknowledge the existence of value-presuppositions implicit in science. This is dishonest. (3) Easlea, Schumacher and others do acknowledge the existence of such value-presuppositions implicit in science. This is more honest still. These writers fail, however, to emphasize the crucial importance of articulating and critically developing such presuppositions as an integral part of intellectual inquiry – thus failing to advocate a rational, critical fundamentalist version of the philosophy of wisdom. In addition, merely to acknowledge that value-presuppositions are implicit in intellectual inquiry is to fail to acknowledge that intellectual inquiry is itself a part of personal, social life, a kind of personal, social action, pursued in order to realize personal, social goals. This is dishonest. (4) Radical Marxists go further, in that they do conceive of, and pursue, intellectual inquiry as an aspect of personal, social action, designed to help achieve personal, social objectives. According to these writers, in capitalist societies

Specialism, then, quite generally, must be rejected. All specialized problem-solving dissociated from fundamental problem-solving must be held to be seriously irrational.

This simple point has profound and far-reaching implications for the whole of scientific, academic inquiry, and for education. For we have seen that scientific, academic inquiry is on the whole at present organized, institutionalized, along specialist, rather than critical fundamentalist, lines. The urgently needed enterprise of discussing fundamental problems in an informal, informed, critical manner – in a manner capable of influencing, and being influenced

intellectual inquiry is devoted primarily to helping to attain the objectives of capitalism: in their intellectual work these writers seek to act in such a way as to help overthrow capitalism, thus creating a socialist society and a socialist intellectual inquiry devoted to helping to realize socialist goals. In so far as these writers see and pursue intellectual inquiry as an aspect of life, social reality, social action, their vision and practice is even more honest still. These writers fail, however, to acknowledge the crucial importance of articulating and critically developing basic socialist presuppositions and objectives. They fail to confront obvious major problems inherent in the idea of a socialist society – such as the problem of centralized, bureaucratic power. This is dangerously dishonest. In particular, as a result of this failing, these writers fail to emphasize the fundamental importance of seeking to develop ways of life, institutions, societies, which progressively develop the aims and methods of personal, institutional and social life – thus enhancing our capacity to achieve that which is of value in life. These writers presuppose answers to problems (3) and (4), instead of seeking to develop a fundamentalist, rational society which enables us to discover improved answers to these problems, as we live.

In short, despite their diversity, the four positions just outlined have one crucial failing in common: they all fail to emphasize that rational action involves quite essentially seeking to improve our aims and methods as we act – the key tenet of aim-oriented rationality (see my 1976a and 1984 or 2007).

by, specialized problem-solving – is obstructed by the prevalence of irrational specialist intellectual standards.

VI

If universalism, and not specialism, provides us with a rational conception of intellectual inquiry, why is it that it is specialism which exercises the predominant influence over most actual scientific, academic work?

The question becomes all the more poignant when we realize how little is new or original in the critique of specialism offered here. Writing over *seventy* years ago now, Aldous Huxley said:

Artistic creation and scientific research may be, and constantly are, used as devices for escaping from the responsibilities of life. They are proclaimed to be ends absolutely good in themselves – ends so admirable that those who pursue them are excused from bothering about anything else. This is particularly true of contemporary science. The mass of accumulated knowledge is so great that it is now impossible for any individual to have a thorough grasp of more than one small field of study. Meanwhile, no attempt is made to produce a comprehensive synthesis of the general results of scientific research. Our universities possess no chair of synthesis. All endowments, moreover, go to special subjects – and almost always to subjects which have no need of further endowment, such as physics, chemistry and mechanics. In our institutions of higher learning about ten times as much is spent on the natural sciences as on the sciences of man. All our efforts are directed, as usual, to producing improved means to unimproved ends. Meanwhile intensive specialization tends to reduce each branch of science to a condition almost approaching meaninglessness. There are many men of science who are actually proud of this state of things. Specialized meaninglessness has come to be regarded, in certain circles, as a kind of hall-mark of true science. Those who attempt to relate the small particular results of specialization with human life as a whole and its relation to the universe at large are accused of being bad scientists,

charlatans, self-advertisers. The people who make such accusations do so, of course, because they do not wish to take any responsibility for anything, but merely to retire to their cloistered laboratories, and there amuse themselves by performing delightfully interesting researches. Science and art are only too often a superior kind of dope, possessing this advantage over booze and morphia: that they can be indulged in with a good conscience and with the conviction that, in the process of indulging, one is leading the 'higher life'.⁶⁷

In fairness to Huxley – in order to excuse the mildness of his words here – we must remember how long ago this passage was written. Since that time, before World War II, everything that Huxley speaks of has of course become much worse.

How and why has this happened? In fact, of course, anyone who has sought to put universalism into practice, and who has explored specialized problems for the light they throw on fundamental problems, will have no difficulty in answering this question. Here, briefly, are some factors responsible for the ever-increasing tyranny of specialism.

1. We fail to put universalism into practice primarily because, as Huxley points out, we fail to take up a measure of personal responsibility for the world in which we find ourselves. And we fail to take up such personal responsibility because of the enormous difficulties that we must inevitably encounter at present in seeking to do so.

These difficulties have arisen as a kind of unforeseen side-effect of the way in which our human world has evolved throughout recorded history. For consider the way in which the problem arises for those who live in the kind of 'human world' experienced by people in pre-historical times – small, closely knit hunting and gathering tribes. In such circumstances, the difficulties that we experience in attempting to assume some personal responsibility for our world do not really arise. Adults – and even children – can without great difficulty assume some measure of personal responsibility for the welfare of the tribe as a whole. All the members of the tribe are known to each individual

⁶⁷ Huxley (1938), pp. 276-7.

personally. Relationships of mutual interdependence are experienced daily, on a personal basis, in hunting, gathering food, and so on. Obligations, responsibilities; towards the tribe can be experienced in a personal, emotional way, in terms of known individuals, in much the same way as we can experience responsibilities towards our family today. (Perhaps the modern, family should be understood as a contraction of the pre-historical tribe.) All members of the tribe have a common outlook on things, a common cosmology and system of values. Thus barriers to intimacy, to mutual understanding, do not arise as a result of differences of outlook and values. Individuals do not face agonizing problems of deciding who they are, how they should live, what there is to give meaning and value to life. On the contrary, the meaning and value of life as lived by the tribe is assured, and is even beyond question, in that no alternative is conceivable. Finally, because of the relative smallness of the tribe, each individual makes a personal impact on the life of the tribe as a whole, and can be well aware of this impact. The tribe, as it were, acknowledges the existence, value and potency of the individual, and is clearly affected by the actions of the individual.⁶⁸

Time passes; agriculture is invented; societies become bigger, more complex, specialized and diversified, requiring much more elaborate, fixed organization. Inter-tribal trade develops: tribes coalesce. Modern methods of travel, transport, and communications develop, and as a result our tribe has become the whole human world, humanity, even, perhaps, life on earth in general.

As a result of these historical developments, the task of assuming some personal responsibility for our common human world has been transformed utterly, and has become almost inconceivably more difficult. Our task is not only to take on some responsibility for the welfare of those who are known intimately to us: rather, in addition, our task is to assume some responsibility – at least to some extent – for the welfare of millions upon millions of complete strangers. No doubt our own welfare is closely bound up with the lives, actions, and welfare of many of these millions of strangers through international relationships such as trade: such relationships of mutual

⁶⁸ For a fascinating account of such a hunting and gathering tribal life, see Turnbull (1976).

interdependence are, however, remote, abstract, not experienced daily on a personal basis. We cannot conceivably experience direct, emotional ties with these millions upon millions of strangers as we do with our friends and members of our own family. Millions of our fellow human beings live lives, see the world, and have values in many ways very different from our own. Not only does this create barriers to mutual sympathy and understanding: responsible concern to understand others – to enter into their different worlds – must inevitably lead us to question the basic assumptions, practices, and values of our own world. The immense diversity of ways of life, cultures, social systems, views of the world, and systems of values with which we are confronted in considering our common human life on this planet must inevitably, at some level, plunge us into doubt and indecisiveness about how to live, what to choose, what to believe and value. And finally, when put into the context of the whole human world, our own life and actions must inevitably, and quite properly, seem to shrink almost to a vanishing point. Unless we possess quite exceptional personal power or influence – something that is perhaps inherently undesirable – all that we do with our lives will have almost no kind of impact or effect whatsoever on the human world as a whole. From this standpoint we are, individually, insignificant and impotent – which may not exactly encourage us to conceive of our world and ourselves from such a standpoint.

For all these reasons it is extraordinarily difficult for the individual today to assume some personal responsibility for our common world. In earlier times this failure did not perhaps matter so much since our power to bring about world-wide changes was strictly limited. Quite suddenly, however, we have developed the capacity to make drastic changes to our whole world. As a result, our common evasion of responsibility has become extremely dangerous for us all. Disasters result. World-wide war, starvation of millions, immense imbalances of wealth and power on a world-wide basis, the population explosion, reckless squandering of irreplaceable natural resources, international politics conducted like gang warfare, the widespread existence of brutal dictatorships, criminal psychopaths (like Hitler, Papadoc, and Amin) even seizing and holding power, the world-wide accumulation of armaments, the constant threat of the nuclear

holocaust, and above all global warming – all these familiar world-wide dangers and disasters are the direct outcome of our general failure to assume personal, adult responsibility for our world.

The members of a small tribal society can, without great difficulty, confront and tackle common problems of the tribe, in a cooperative, responsible fashion. Tribal meetings can be convened at which everybody can be free to articulate problems, propose and criticize possible solutions.

In our modern world this cannot be done. The population of the earth cannot hold a meeting to discuss common problems where everyone is free to speak. And yet something like this must exist if general understanding of, and responsibility towards, our common human problems is to develop at the personal level – something that we must develop if we are to be able to cope with the dangers and disasters just indicated. We cannot rely on existing institutions, existing centres of power, existing governments, whether democratic or dictatorial: all this is all too blatantly failing at present to cope adequately, i.e. humanely and rationally, with our problems. In the end the point is very simple. In the absence of general understanding of, and responsibility towards, our problems, genuinely democratic governments responsive to public opinion will be unable to act responsibly as far as our most urgent, general, common problems are concerned.⁶⁹ Public opinion will not permit it. In a sense, only undemocratic, dictatorial governments, capable of suppressing or ignoring public opinion, will be able to act in such a fashion. Dictatorships, however, put us at the mercy of the decisions and actions of those few individuals who have won the fight for power (thus being, almost inevitably, ruthless and power-mad). Either way it is most unlikely that global problems will be tackled responsibly. For this we need a widespread, even world-wide understanding of, and responsible attitude towards, our basic problems at the personal level. And for this in turn it is essential that we develop a modern, world-wide institutional equivalent of the tribal meeting.

⁶⁹ Higgins writes (1978, pp. 21-45) especially clearly and convincingly on this point, in part from personal experience.

It is in this way, I suggest, that we need to conceive intellectual inquiry – as the open, sustained, responsible discussion of our common problems. Intellectual inquiry needs to be conceived and pursued as the tribal meeting of humanity, permanently in session, open to all, our joint endeavour to develop cooperative, personal responsibility for our common problems. *Something* must be created to replace the tribal meeting. Intellectual inquiry, at its best, constitutes such a replacement: it is from this standpoint that intellectual inquiry needs to be understood, contributed to, and judged.

And only universalism can do justice to this conception of intellectual inquiry. This, indeed, *is* universalism: intellectual inquiry conceived as the outcome of our personal, cooperative, responsible attempts to improve our solutions to our fundamental problems.⁷⁰

The difficulty we experience, then, in putting universalism into practice is an important part of the difficulties we experience in seeking to take on a degree of personal responsibility for our shared world. Specialism is, as Huxley correctly points out, an evasion of responsibility, the outcome of a failure to cope with the stress of responsibility. Specialism can even be seen as the outcome of a

⁷⁰ A humane, cooperative, mutually understanding, pluralistic society presupposes and is, in a sense, presupposed by, universalism. If two people, two societies, or two cultures, giving different answers to our four fundamental problems, are to act humanely and cooperatively together, there must be mutual understanding; this requires that each is able to imagine, at least as a possibility, that the other's answers are correct (or at least an improvement, in certain respects, over his own). This in turn requires that each recognizes the genuineness of the four fundamental problems. If each is to learn from the other, then each must acknowledge the genuineness of the four fundamental problems. On the other hand, to recognize that these problems are genuine is to imagine at least the possibility of answers different from one's own being given – which is to imagine a pluralistic society, at least as a possibility.

Only universalism can do justice to the Socratic and Kantian idea that Reason forms a basis for the unity of mankind.

kind of intellectual or professional tribalism – the specialist's tribe being the 'invisible college' of like-minded specialists.

A number of writers have been concerned to emphasize – in terms somewhat analogous to those outlined here – that the blessings resulting from moving from the intimate, coherent tribe to the big, complex, diversified modern world are mixed. These writers all emphasize, in one way or another, that this transition makes possible the development of choice, freedom, reason, science, on the one hand, but can also lead to uncertainty, fear, loneliness, a sense of meaninglessness and impotence, on the other hand.

Thus in *Coming of Age in Samoa*⁷¹ Margaret Mead tells us that children in Samoa fail to experience anything like the trauma of adolescence so familiar in Western society. She concludes that this is due to the absence in Samoa of the problem of choosing between rival ways of life and values. Adolescent trauma, then, is due to the great difficulties that we experience in coming to terms with cultural diversity in our society – in turn due, without doubt, to a general failure of our culture to cope adequately with this central problem of diversity. As I have already indicated, in *The Open Society and its Enemies* Popper argues that the open society – the society in which diverse ways of life are tolerated – is essential to our humanity, our reason, our civilization. It is only with the existence of social diversity that we can begin to doubt, to criticize, to learn, and perhaps to make progress. In Popper's view, rationality is to be understood primarily in terms of the capacity to doubt, to criticize, and thus to learn: criticism, however, is only really possible if a plurality of views and ways of life coexist in society. Thus, for Popper, rationality is to be understood primarily in social terms, arising as a result of social developments – the development of social and cultural diversity, and a tradition of criticism.⁷² The development of the open society

⁷¹ Mead (1943).

⁷² 'Reason, like language, can be said to be a product of social life' (Popper, 1966a, vol. 2, p. 225). See also the discussion of the claim that science is necessarily social in character, 'Robinson Crusoe science', however successful, being necessarily only 'revealed science' in that it must lack objectivity, pluralistic criticism (ibid., pp. 216-20). Unfortunately, Popper in his later work fails

lamentably to develop these anticipations of the point stressed in this paper (see, e.g., note 51), that reason, epistemology, thought, intellectual inquiry, all need to be conceived of, and developed, as personal and social in character, in the world, a part of life.

If we adopt the view advocated in section VII that the aim of intellectual inquiry is to help us achieve wisdom, life of value, then the fundamental aim of intellectual inquiry becomes a personal, social aim, and the problems of intellectual inquiry become, fundamentally, personal, social problems of living. Our central task, in pursuing intellectual inquiry, becomes to help develop more rational, wiser ways of living, institutions, social orders. The split between personal, social aims and intellectual aims – the split between personal, social action and thought – disappears. Popper, however, holds that the basic aim of intellectual inquiry is to develop impersonal, objective knowledge. This leads him to develop his 'world 3' theory of the intellectual domain. As a result, and quite disastrously, the fundamental personal and social problems of intellectual inquiry – problems we encounter in helping to develop life of value, a wiser world – are transformed into the pseudo metaphysical-neurological problems of how 'world 3' can interact with the mind and the brain: see Popper (1972); Popper and Eccles (1977).

Universalism, in sharp disagreement with Popper, recognizes just one world. Within universalism, Popper's conceptually incoherent psycho-neurological thesis that world 3 interacts with world 1 via world 2 can be replaced by the kind of conceptually coherent psycho-neurological postulate indicated in Note 8, or by a version of this postulate which asserts that *aim-oriented rationalism* is programmed into the neurological structure of our brains – or at least needs to be so 'programmed' if we are to be able to achieve what is of value in life. (For an exposition of aim-oriented rationalism see chs. 7 and 8 below and my (1976a, chs. 8 and 9; 1984 or 2007, ch. 5). There is only one world; it is in *this* world that universalism and aim-oriented rationalism need to flourish. In order to help achieve this, it is essential that we see critical fundamentalist intellectual standards, aim-oriented rationality, as something embedded, actually and potentially, in this

makes possible the development of both freedom and reason. Popper is at pains to emphasize, however, the price we pay for these developments – the strain that civilization puts upon us. It is indeed the major thesis of *The Open Society and its Enemies* that the uncertainties, the emotional stress, created by our movement towards the open society can be so great that we long passionately for a return to the simplicities and certainties of the monolithic closed society. This anti-rational, anti-humanitarian longing is responsible for the totalitarianism of both left and right. The difficulties that confront us in coming to terms with the open society are indeed, according to Popper, so extreme that even many of our greatest thinkers in the past have failed to surmount them: Heraclitus, Plato, Aristotle, Hegel, and Marx all in one way or another, in Popper's view, sought to return us to the closed society. Many of our greatest philosophers and rationalists have been *enemies* of the open society.⁷³

world. It is essentially this insight that we need to see implicit in much of Popper's *The Open Society and its Enemies* in order to appreciate the real value of that work, and in order to make sense of Popper's wonderful suggestion that we should see intellectual evolution as a development of biological evolution. In the circumstances it is somewhat tragic that Popper should have gone on, with the development of his three-world view, to reject explicitly the insight that reason needs to be seen as materially and socially embodied in this one world.

⁷³ According to Popper, we must learn to live with the intense emotional strain of civilization, as the price that must be paid for reason, for the open society, for civilization. Any attempt to introduce social and cultural changes which alleviate this strain must be fiercely resisted, as such changes must inevitably lead to totalitarianism. However, as indicated in section II, this is because Popper fails to conceive of the possibility of universalism, and is led as a result to defend a seriously irrational and undesirable conception of rational inquiry and civilization.

For purposes of clarification, I should indicate four further main differences between the viewpoint being advocated in this essay, and views advocated or presupposed in Popper's writings. The chief difference, unquestionably, is simply this. I advocate that the

basic aim of rational inquiry is to enhance wisdom. Here I part company not only with Popper, but with the whole Western tradition, in that this tradition gives to rational inquiry the basic aim of enhancing *knowledge* (human welfare, enlightenment, and progress being only secondary and uncertain by-products of the basic and prior achievement of knowledge). As a result of giving priority to wisdom – to our living, actual capacity to discover and achieve what is of value in life – I am led to locate rational inquiry, at the most fundamental level, within and amidst our lives, personal and interpersonal or social. This leads me to stress the fundamental importance of *aim-oriented rationalism* designed to help us improve our aims, and thus our lives, as we live (*aim-oriented empiricism* being simply a special case of aim-oriented rationalism, applicable to science). This is in marked contrast to Popper's conception of reason, which he has called *critical rationalism* (*falsificationism* being simply, for Popper, a special case of critical rationalism, applicable to science). Since for me rational inquiry has, as its basic task, to help us achieve what is of value in life, I hold that all intellectual values need ultimately to be founded in human value – especially in the supreme value of each individual person, and the good things that can go on between people once this is recognized – rigorous, objective inquiry being, as though by definition, a universal tool perfectly designed to help us all maximally to achieve, or grow, life of value. I thus disagree absolutely with Popper's thesis – in effect a standard component of the philosophy of knowledge – that purely scientific values should be distinguished as sharply as possible from human or extra-scientific values. (This is Popper's *sixteenth thesis* in his 1976b, pp. 96-98.) The purpose of the present essay is to argue that the rationality, the intellectual rigour, the objectivity, of inquiry is essentially bound up with the capacity of inquiry to help us resolve those problems of living we need to resolve in order to achieve what is of value in life. That which is of value in life is primary: intellectual value is a reflection of – or is subservient to – primary value in life. Ideas – including the idea of this essay – are spectacles intended to help us to see clearly what is of value in existence, actually and potentially; they are forks and spades

designed to help us to cultivate what is of value in our lives, in reality. Like spectacles, ideas are to be assessed in terms of whether they serve to clarify or blur our vision; like forks and spades, ideas are to be assessed in terms of their use, their success in practice. The idea that intellectual value is dissociated from value in life, in the world, quite fundamentally misconceives the proper value and use of ideas: pushed to the extreme this becomes Plato's doctrine of the forms. (The Popperian, Western doctrine of the autonomy of intellectual value devolves, in fact, from Plato's doctrine.) The idea that intellectual value needs to be conceived of as dissociated from value in the world – and not as integral to and contributing to value in the world – receives support no doubt from the desire of many intellectuals to find in intellectual work some kind of escape from the world, a quiet and transparent refuge. Given Popper's defence of the orthodox doctrine concerning the autonomy of purely scientific or intellectual value, it is not at all surprising that he should call Hume's thesis that 'Reason is, and ought only to be, the slave of the passions' a 'horrifying doctrine': Popper (1977, p. 132). An upholder of aim-oriented rationalism and the philosophy of wisdom would wish to make only minor adjustments to Hume's thesis. Either reason should be held to be the slave of that most profound passion of ours of all to participate in life of value: or reason ought perhaps to be called the 'good servant' or 'enlightened tutor' of the passions – reason itself the outcome of our cooperative, balancing, or resolving passion for a whole, authentic life of value. As a result of conceiving of inquiry and reason as being an essential, active component of human life, more or less realized in practice in our personal and social actions, I am led to avoid the conceptual incoherence of Popper's three-world view, as indicated in notes 8 and 53. In general, the viewpoint that I wish to advocate is much closer to Einstein's than to Popper's, taking into account especially the emphasis that Einstein came to place in his later life on the fundamental importance of developing a living ethical culture, and a kind of education designed to help us acquire and participate in such a culture. 'It is not enough to teach a man a speciality. Through it he may become a kind of useful machine but not a harmoniously

It is scarcely surprising, then, that adolescents, emerging from the 'closed' society of the family into our quasi-'open' society, should experience difficulties. The problems of adolescence need to be understood in philosophical or rationalistic terms – in terms of emotional reactions to an intense awareness of possibilities and uncertainties – and not merely in terms of some psychological theory of emotional development.

Isaac Bashevis Singer, in his novels and short stories, has given us a wonderfully vivid and perceptive account of the enormous difficulties we encounter in emerging from a closed society.⁷⁴ In *The Manor* and *The Estate* Singer provides us with a wholly convincing picture of the confusion, the sense of loss, that overwhelmed those enlightened Jews who, towards the end of the last century, emerged from the highly traditional, almost mediaeval, Jewish communities still existing then in Poland. Singer's writings are especially noteworthy for the fact that many of his protagonists are themselves deeply conscious of the problem, and not merely affected emotionally by it without any understanding of its nature. Singer is concerned to show us, in a fictional form, individuals grappling passionately with the task of pursuing critical fundamentalist intellectual inquiry. Singer's vivid and honest imagination takes us to the heart of the problems of our civilization.

Essentially the same problems – explored by both Popper and Singer – have also been discussed by Erich Fromm, for example

developed personality. It is essential that the student acquire an understanding of and a lively feeling for values. He must acquire a vivid sense of the beautiful and of the morally good. Otherwise he – with his specialized knowledge – more closely resembles a well-trained dog than a harmoniously developed person. He must learn to understand the motives of human beings, their illusions, and their sufferings in order to acquire a proper relationship to individual fellow-men and to the community': Einstein, 'Education for Independent Thought', in his (1973, p. 66).

⁷⁴ See Isaac Bashevis Singer (1975a; 1975b; 1975c; 1977a; 1977b; 1977c).

in his *The Fear of Freedom*.⁷⁵ Finally, Peter Gay, in his marvellous book *The Enlightenment: An Interpretation*, provides us with a haunting account of the anguish experienced by the thinkers of the Enlightenment in attempting to come to terms with their doubts, their scepticism, as they emerged from the religious tribalism of contemporary Christianity.⁷⁶

It is, I hope, clear that all these writers are concerned essentially with the same problem – the difficulties we encounter in coming to terms with something that is essentially desirable, namely social and cultural diversity. One disastrous consequence of specialism is that it disrupts understanding of problems as fundamental as this: the problem is scattered amongst a number of disparate disciplines, and lost sight of. Instead of discussion being organized around the problem, so that contributions such as those of the above writers can fruitfully interact with, and supplement, each other, discussion is organized instead within the disciplines: anthropology, epistemology, political philosophy, history, psychology, history of ideas, fiction. As a result, we fail to discover the interconnections between the contributions: we fail to improve our understanding of the underlying problem. We fail to understand the problems of adolescence as those of moving from a closed to an open society – in part philosophical problems. We fail to appreciate the social, cultural, and personal implications of Popper's philosophical and epistemological discussions. We fail to grasp the universal significance of Singer's fiction. We do not see that Popper and Singer are concerned with essentially the same problem. Fromm may be dismissed as pursuing the pseudo-scientific discipline of socio-cultural-psychoanalytic psychology, instead of being understood as contributing to our understanding of the problems discussed in *The Open Society and its Enemies*, and in Feyerabend's 'The Problems of Empiricism'.⁷⁷

⁷⁵ E. Fromm, *The Fear of Freedom*, Routledge & Kegan Paul, London 1960. See also, for example, his *The Sane Society*, Routledge & Kegan Paul, London 1963.

⁷⁶ Gay (1973, vol. 1, pp. 59-71).

⁷⁷ Feyerabend (1965).

Specialism thus prevents us *from seeing* our fundamental problems. As a result, we fail to see the urgent need to improve our thinking at this level, and the considerable difficulties that arise in connection with this task.

2. It is in the nature of universalism to raise questions and doubts that can be highly awkward for those who wield power in society. In particular, of course, universalism challenges all those who claim to have authoritative answers to fundamental problems – religious and secular centres of power and influence in society. Universalism calls into question cherished beliefs and values, and thus also is liable to collide with public opinion. Powerful social forces, then, will inevitably discourage the development of critical fundamentalist intellectual inquiry – as Socrates, Galileo, and Spinoza, for example, found out. Only a society which had, quite generally, taken universalism to heart would encourage the development of critical fundamentalist intellectual inquiry: but of course no such society has as yet come into existence.

The case of specialism is, however, quite different. Specialist scientists and scholars may well be quite content to let non-academic authorities decide fundamental issues, scientific, academic inquiry confining itself to solving those specialized, technical problems whose solutions are required by those who wield power in society. Critical fundamentalist issues in any case lie beyond the reach of specialist intellectual standards and concerns. Specialism thus robs the scientist and scholar of the capacity, from a professional standpoint, to criticize fundamentalist decisions made on the basis of power in society – except where those in authority are foolish enough to transgress specialist standards and results.

3. Specialism is especially appealing to those who uphold what may be called 'oracular' conceptions of reason – according to which reason, ideally, is something that reaches decisions for us, rather than being something which helps us to decide. In terms of

such oracular conceptions of reason, universalism is, of course, indefensible.⁷⁸

But all such oracular conceptions of reason must be rejected. We may identify reason with some set of rules, laws, methods, or criteria which dictate decisions to us. It ought always to be, however, our own decision to adopt these laws, methods, etc. Genuine rationality involves being able to choose and develop such laws to suit our purposes. Universalism is correct in insisting that genuine rationality involves recognizing that ultimately *we* choose and decide.

4. According to specialism, the expert is entirely entitled to pronounce authoritatively on matters relating to his discipline – in a manner which ignores the contributions, the criticisms, of non-experts. This is because, according to specialism, only specialized considerations can be relevant for an assessment of specialized results. Only the expert can be competent to contribute to a specialized discipline. There can be no doubt that being able to pronounce authoritatively in this kind of way is something that is deeply appealing to many. Universalism, however, deprives the expert of this deeply appealing authoritative immunity from outside criticism. Basic assumptions about the nature of the world, and about the meaning and value of life, must inevitably, according to universalism, pervade specialist work. It cannot be correct for experts to decide for the rest of us what these assumptions should be. It is thus entirely proper that non-experts should be able to challenge and contribute to critical fundamentalist assumptions implicit in specialized work. It is indeed important that experts do listen to non-expert comments and criticisms concerning fundamental assumptions, since it is all too easy for the expert to forget the prevalence and influence of such assumptions amidst his technical work – losing sight of the wood for the trees.

⁷⁸ It is doubtless commitment to this kind of oracular conception of reason which leads both Kuhn and Feyerabend, in their rather different ways, to characterize science as irrational. The same mistake is implicit in almost all forms of relativism.

5. Increasingly, during the last fifty to one-hundred years, scientific, academic work has become something that is engaged in as a profession, a career, rather than out of amateur love. The scientific, academic enterprise has become increasingly institutional and bureaucratic in character. All this favours, and almost requires, specialism. For these factors require that scientific, academic work can be assessed in a definite, agreed way, sound work being distinguishable from unsound work in an uncontroversial manner. Promotions, funding of research work, professional status, management of research – these career and institutional matters all favour the adoption of definite, agreed specialist intellectual standards. Sustained inquiry into fundamental problems is much more difficult to professionalize and institutionalize. Crucial institutional questions such as whose work is to receive funds, be taught, be rewarded with promotion and academic honours, become almost impossible to decide in a standard, bureaucratic manner.⁷⁹

6. Once a conception of intellectual inquiry has become established – built into the institutional and bureaucratic structure of intellectual inquiry – all sorts of mechanisms tend to preserve this institutionalized conception. Education will tend to indoctrinate pupils and students in this conception. Only those who conform to the standards of the conception will be able to do research work, publish, obtain academic jobs. Only that work which conforms to the accepted standards will be published, and will be accepted on publication. Even those who disagree with the institutionalized viewpoint will be

⁷⁹ In this paper I am of course arguing that even though the goal is difficult to attain, nevertheless it is a matter of supreme importance that we seek to build universalism, rather than specialism, into the institutional structure of the scientific, academic enterprise, and education. Indeed what I am proposing goes much further than this. We need to build universalism, and aim-oriented rationalism, into our whole way of life, into society as a whole, into the human world. The basic aim of intellectual inquiry ought to be to devote reason to the enhancement of wisdom. This programme can scarcely begin to be put into practice, however, as long as specialism rather than universalism is built into the institutional structure of the scientific, academic enterprise.

obliged to pay lip service to it, simply in order to teach, publish, and do research. As a result, the public face of scientific, academic inquiry will come overwhelmingly to conform to the general viewpoint, and it will seem increasingly absurd to call this viewpoint into question.

Once specialism is established institutionally, in short, no problem arises as to why this viewpoint should persist.

These, then, are some of the factors responsible for the failure to put universalism into practice – responsible for a pervasive corruption of intellectual standards.

VII

If universalism were to be put into practice we would expect intellectual inquiry as a whole to give priority to our most general and important problems – specialized problems being chosen and tackled in order to help us solve the former.

The result to be expected from putting specialism into practice is, however, the exact opposite. Although specialized, technical problems may well be tackled with brilliance and great success, from the standpoint of what matters most in life the vast industry of specialized problem-solving may well seem largely irrelevant. Most specialized problem-solving will be unrelated to our fundamental problems. Specialized problems will not be understood or tackled as subordinate problems to fundamental problems. Instead of illuminating our understanding of how fundamental problems may be solved, intellectual inquiry will tend to do the exact opposite. We will tend to be overwhelmed by a vast maze of specialized disciplines, jargon, and results. It is not just intellectual inquiry as a whole that will suffer as a result. *We* will suffer. Our capacity to think and act intelligently, in response to our basic problems, will be sabotaged. Experts will become, not our servants, but our masters.

In his Nobel peace prize lecture, Martin Luther King declared:

Modern man has brought this whole world to an awe-inspiring threshold of the future. **H**e has reached new and astonishing peaks of scientific success. He has produced machines that think and instruments that peer into the unfathomable ranges of interstellar space. He has built gigantic bridges to span the seas

and gargantuan buildings to kiss the skies. His airplanes and spaceships have dwarfed distance, placed time in chains, and carved highways through the stratosphere. This is a dazzling picture of modern man's scientific and technological progress.

Yet, in spite of these spectacular strides in science and technology, and still unlimited ones to come, something basic is missing. There is a sort of poverty of the spirit which stands in glaring contrast to our scientific and technological abundance. The richer we have become materially, the poorer we become morally and spiritually. We have learned to fly the air like birds and swim the sea like fish, but we have not learned the simple art of living together as brothers.⁸⁰

The predominance of specialist intellectual inquiry plays its part, I suggest, in the development of the 'glaring contrast' to which Martin Luther King here refers – the achievement of specialist knowledge at the expense of the achievement of wisdom.

Consider the following analogy. Our problem, let us suppose, is to build a house. On the one hand we may tackle this problem in a critical fundamentalist manner. We propose and criticize possible solutions to our basic problem – thus developing an over-all plan. In order to solve our basic problem, however, a host of specialized, technical, subordinate problems need to be solved. Bricks need to be made; so, too, slates, doors, window frames, windows, beams, plaster, floorboards, and so on. Foundations need to be dug and cemented. All the various parts need to be assembled properly, in conformity with the plan, to build the house. Plumbers need to put in pipes, tanks, sinks, a bath; electricians need to wire the house; and so on. An intricate maze of highly specialized, technical problems need to be solved by an army of experts if the house is to be built. Equally, however, if the house is to be built, it is absolutely essential that the specialized problem-solving be properly coordinated so that it all gives rise to a solution to the fundamental problem – to build a house. There needs to be a constant two-way flow of information between problem-solving at the fundamental

⁸⁰ Haberman (1972, pp. 333-4).

level, and at the specialized level. Failure to solve certain specialized problems may necessitate a revision of the basic plan.

This common-sense, critical fundamentalist approach is in complete contrast to a specialist approach. According to specialism, building a house only involves solving specialized, technical problems. The fundamental problem – what kind of house do we want? – is not a problem that the building trade can take seriously. (It is meaningless, subjective, incapable of being decisively solved, philosophical, evaluative or religious; in any case not a matter for the trade to concern itself with professionally.) The building trade needs to concern itself with specialized, technical 'puzzles' – manufacturing bricks, mortar, cement, slates, floorboards, windows, wiring, pipes, and so on. Progress in the building trade is to be judged in terms of how well these specialized puzzles are being solved.

The outcome of all this will of course be ever increasing piles of completely unusable bricks, slates, wire, pipes, etc. – and no house will be built at all. And if we complain, we will no doubt be met with indignation in that each specialist has indeed performed his task with skill and expertise.

The 'house' that intellectual inquiry as a whole should help us build is, I suggest, a life of value – a rich and fulfilling life, a life in which we can share friendship, love, happiness, beauty, creative work, joy in being alive. Our 'fundamental' problems are the problems we encounter in our lives in seeking to discover, experience, participate in, help create that which is of value. The basic rationale for the whole of intellectual inquiry is to help us to articulate and solve these fundamental problems of living. All intellectual problems are subordinate to these fundamental personal and interpersonal problems of living. The problems of mathematics, logic, philosophy, theoretical physics, cosmology, molecular biology, neurology – all these need to be understood as sub-problems of our fundamental personal and interpersonal problems of living.

It is, of course, not the case that intellectual inquiry is pursued only for pragmatic reasons – as a means to the realization of non-intellectual, practical ends. Intellectual inquiry is also pursued for its own sake. Intellectual inquiry is, in other words, itself a part of life, enriching life directly when pursued for its own sake, like music or poetry. It is, for example, of the essence of life of value that we are

perceptive and curious about our surroundings – in touch with our environment. 'Pure' research in physics, say, or cosmology, geology, history or anthropology, amounts simply to a cooperative following-up of such personal perceptiveness and curiosity. From the standpoint of pure intellectual inquiry, it is the curiosity, the imaginative explorations, the thoughts and feelings, the knowledge and understanding, the intellectual honesty and passion, the problem-solving, of people in society as a part of life, that really matters. It is our shared exploration of our world, as an aspect of life of value, that is important. It is the personal knowledge and understanding of our world that we have ourselves developed, integrated into our lives, that really matters. Pure intellectual inquiry is, in other words, at the most fundamental level, personal and interpersonal in character, a part of life. The impersonal or institutional aspects of pure intellectual inquiry exist simply as a means to an end: to aid personal and interpersonal curiosity, wonder, knowledge, and understanding, as a vital aspect of our personal and social lives. Thus both 'pure' and 'pragmatic' intellectual inquiry seek to contribute to the richness, the value, of our shared lives here on earth. In both cases, what ultimately matters is the value of our personal and interpersonal lives.

The fundamental aim of intellectual inquiry, we may say, is to enhance our personal and interpersonal wisdom – our capacity to discover and achieve what is of value in life both for ourselves and for others. All intellectual problems are problems subordinate to our basic life-problems of wisdom. Of the four fundamental problems formulated above, it is the third and fourth that are the most fundamental, the first and second being pursued as a part of our concern to discover and achieve what is of value in life.

If intellectual inquiry is to meet with success in helping us to discover and achieve what is of value in life, then it must of course be generally understood to have this basic purpose. Education must enable us to come to understand and use intellectual inquiry in this kind of way,⁸¹ so that we discover fruitful interconnections between

⁸¹ See Gray (1972) for the suggestion that the aim of education should be the achievement of wisdom in life. Gray fails to point out, however, that intellectual inquiry is at present profoundly

our personal problems and 'impersonal', 'objective' intellectual problems, our own personal, childish wonderings about the nature of the universe, for example, illuminating and being illuminated by the 'official' wonderings of Kepler, Newton, Faraday, or Einstein, or our personal problems of adolescence illuminating and being illuminated by the philosophical, social problems of the open society discussed, for example, by Popper, Fromm, Mead, and Singer. Intellectual inquiry must itself be organized in such a way as to be amenable to this kind of understanding and use. Above all, scientists and scholars must be fundamentally concerned to develop intellectual inquiry in such a way that it is designed to help us build our 'houses' of wisdom with our lives. All this is essential if intellectual inquiry is to be developed as the tribal discussion of humanity, designed to help us create more valuable lives, a better human world.

When viewed from this perspective of the philosophy of wisdom, present-day scientific, academic inquiry is, in terms of our analogy, more like an unusable, chaotic heap of bricks, slates, window frames, and pipes, than something out of which we can build a habitable house. What confronts us is an immense pile of specialized jargon-ridden disciplines pursuing specialized intellectual problems dissociated from our problems of living, there being little indication as to how the non-specialist is to find his way through all this to discover and achieve what is of most value in life. Scientific, academic inquiry is not pursued, understood, taught, or organized in accordance with the over-all assumption that what ultimately matters is personal and social wisdom.

In so far as a basic organizing assumption is built into present-day scientific, academic inquiry, it is that the aim of such inquiry is to improve objective, impersonal, institutional *knowledge*, not personal and social *wisdom*. Intellectual priority is not given to our problems of living, to the difficulties, frustrations, sufferings that we encounter in our lives in attempting to discover and achieve what is of value in life: on the contrary, intellectual priority is given to impersonal problems of knowledge encountered by the various academic disciplines in seeking to describe, predict, and explain phenomena.

irrational and defective when judged from the standpoint of having as its basic aim to help us achieve wisdom.

Even the social sciences give intellectual priority to problems of knowledge as they arise within sociology, psychology, and so on, rather than to the problems encountered by people in their lives. Intellectual progress is assessed, not in terms of the success we meet with in achieving what is of value in life, but rather in terms of the success achieved in acquiring academic knowledge. Intellectual progress is conceived as being decisively dissociated from human, social progress.

Impersonal, academic problems of knowledge may of course be tackled out of a concern to develop knowledge which can subsequently be used or applied to help solve human, social problems. The all-important point, however, is that these problems of knowledge are neither understood nor tackled as intellectually subordinate to our more fundamental problems of living, but are, on the contrary, decisively dissociated from these. If science is to be of human value, it tends to be argued, it is essential that science acquires reliable, objective, impersonal factual knowledge, this in turn requiring – so the argument goes – that the problems of knowledge be tackled in a way which is decisively dissociated from the problems of life.

From the standpoint of developing a kind of intellectual inquiry designed to help us achieve what is of value in life, however, all this is irrational in a quite elementary fashion, and for precisely the reasons emphasized throughout this essay. Granted that the fundamental task of intellectual inquiry is to help us solve those personal, social problems of living we encounter in seeking to achieve what is of value in life, elementary rules of rational problem-solving require us to give intellectual priority to the task of articulating these personal, social problems of living, and proposing and criticizing possible (and actual) solutions to them. Rationality also requires, of course, that we develop a multitude of subordinate, specialized problems – for example, technological problems, scientific problems, problems of knowledge and understanding. It is absolutely essential for rationality, however, that these specialized problems are understood as subordinate, the enterprise of tackling them being set within the framework of the more fundamental intellectual activity of proposing and criticizing possible solutions to our problems of living.

The philosophy of knowledge is, as I have said, at present almost universally taken for granted by the academic community, and is built into the whole institutional structure of the scientific, academic enterprise. As a result, the elementary irrationality of this philosophy has damaging repercussions for the whole of intellectual inquiry, and indeed for the whole modern world, all our lives here on earth. Both 'applied' and 'pure' intellectual inquiry, it should be noted, are damaged by the general acceptance of the philosophy of knowledge.

On the one hand we may – with Bacon, Comte, Bernal, and Ravetz, for example – be concerned primarily with the capacity of intellectual inquiry to help us solve our practical social problems. If so, then according to the philosophy of wisdom, intellectual priority needs to be given to articulating these problems, and proposing and criticizing possible solutions. Solutions to practical social problems are appropriate personal, social *actions*. Hence, according to the philosophy of wisdom, the fundamental intellectual task of intellectual inquiry is to develop imaginatively and assess critically possible and actual personal, social *actions*. The development of knowledge and technology needs to be rationally subordinated to the more fundamental intellectual activity of proposing and criticizing social *actions*.

The philosophy of knowledge, however, gives intellectual priority to the development of knowledge divorced from a concern with our social problems. New knowledge leads to the development of new technology which is then applied in ways which help, we may hope, to solve these problems. The crucial point, however, is that intellectual priority is given to the task of proposing and criticizing *claims to knowledge* – laws, theories, experimental results – instead of possible social actions.

Inevitably, as a direct result of giving intellectual priority to the development of knowledge rather than to proposing and criticizing possible solutions to social problems, intellectual inquiry must: (1) fail to help us solve all those major social problems which require new social actions, policies, and institutions for their resolution rather than new knowledge and technology; (2) fail to help us give priority to the development of new knowledge and technology most needed for the resolution of urgent social problems; (3) fail to help us

use such knowledge and technology, where developed, to maximum advantage in a rational fashion to help solve social problems; (4) fail to help us anticipate and prevent new knowledge and technology being used in socially harmful ways; (5) fail to help us anticipate and refrain from engaging in intrinsically harmful scientific research; (6) fail to concentrate intellectual attention on our most urgent social problems.⁸² These six kinds of failure are all immediate consequences of the fundamental failure to give intellectual priority to rational human, social problem-solving. As long as our thinking about the world and ourselves is dominated by the philosophy of knowledge, it is almost inevitable that the social ills of the modern world will arise, even if almost everyone acts with good will.

On the other hand we may – with Kepler, Spinoza, Einstein, and Popper, for example – be concerned primarily with the 'intrinsic' or cultural value of intellectual inquiry, intellectual inquiry pursued for its own sake. If so, then we need to recognize – as emphasized by the philosophy of wisdom – that it is knowledge and understanding achieved by people that ultimately matters. 'Pure' intellectual inquiry, conceived of in impersonal or institutional terms, is of value in so far as it helps us to achieve that which really has value – our personal knowledge and understanding of our world, our personal curiosity, perceptiveness, capacity to discover that which is of significance in our surroundings, and the extent to which all this enriches our life. The problems of 'pure' intellectual inquiry are, in other words, at the most fundamental level, personal and interpersonal problems, problems that we encounter in seeking to enhance our personal knowledge and understanding of the world, our personal perception and appreciation of what is significant and of value in existence. As Einstein once remarked: 'Knowledge exists in two

⁸² Only if intellectual priority is given to the task of proposing and criticizing possible and actual personal, social actions, policies, aims and methods, institutional enterprises, ideologies – problems of knowledge and technology being tackled as subordinate to our fundamental personal, social problems of living – can intellectual inquiry overcome these defects. For an elaboration of this point see my (1984 or 2007; 2010).

forms – lifeless, stored in books, and alive in the consciousness of men. The second form of existence is after all the essential one; the first, indispensable as it may be, occupies only an inferior position.⁸³ To this I would only add that from the standpoint of 'pure' intellectual inquiry it is perhaps the activity, as a part of life, of imaginatively exploring the world, following up our passionate curiosity, the lively encountering of aspects of reality, that is the thing that is essentially of value. And just as the professional, specialized, institutionalized activities associated with music are designed, ideally, to further our making and enjoying of music, so too the professional, specialized, institutionalized activities associated with science are, ideally, designed to further our exploration and enjoyment of our world.

All this is in marked contrast with the views of those who, like Popper and Ziman, emphasize the fundamental importance of 'objective knowledge', of 'knowledge without a knowing subject', of 'public knowledge', or of 'institutional knowledge', conceived as ends in themselves, rather than as means to the achievement of the end of life of value, via enhancement of personal awareness of the world.⁸⁴ In insisting that 'pure' science be dissociated from life, intellectual progress being understood in wholly objective, impersonal, or institutional terms, the philosophy of knowledge misses out precisely that which matters most, our personal apprehension of the world. As a result of putting this philosophy of impersonal knowledge into practice, a disastrous split develops between the way we personally apprehend or conceive of the world, and the way 'science' apprehends or conceives of the world. We fail to exploit science in order to enrich and extend our personal vision of things; and we fail to develop science in such a way that it is amenable to such exploitation. We fail to discover how to use scientific theories as spectacles through which we may, conjecturally, view the world. Instead of emphasizing the priority of the personal problems of understanding we need to solve in order to make such a use of scientific theories, the problems are dismissed as 'subjective', the development of impersonal knowledge embodied in scientific theories becoming an end in itself. As a result we become blind to – or ignore –

⁸³ A. Einstein (1973, p. 80).

⁸⁴ K. Popper (1972); Ziman (1968).

the profound discrepancies that exist between the world as conceived by us in life, and the world as conceived, impersonally, by science. A kind of advanced intellectual schizophrenia in our thinking develops. Theoretical physics, for example, ceases to be, with Einstein, a personal 'attempt conceptually to grasp reality as it is thought independently of its being observed',⁸⁵ and becomes merely the impersonal, institutional, ritualistic prediction of phenomena, 'the whole thing . . . a wretched bungle . . . which can 'only claim the interest of shopkeepers and engineers'.⁸⁶ Personal awareness of what is significant and of value in existence, intellectual passion, curiosity, wonder, all degenerate into nothing more than the possession of information and expert skills, the accumulation of dry knowledge of fact. As a result of dissociating 'pure' intellectual inquiry from life, we lose sight of the value which intellectual inquiry has when pursued for its own sake.

Above all, and quite generally, as a result of engaging in, and thinking in terms of, intellectual inquiry as in the first instance the pursuit of impersonal knowledge we lose sight of those problems which, quite fundamentally, create the need for intellectual inquiry, and which intellectual inquiry ought fundamentally to be helping us to solve. By giving priority to the pursuit of impersonal knowledge, we fail to emphasize the fundamental character of the personal and social problems of our pluralistic world. Intellectual inquiry must then fail to enhance our common understanding of these problems and our common capacity to develop more adequate resolutions to them. Conceiving of things in terms of the pursuit of impersonal knowledge, we fail entirely to see the urgent need to develop intellectual inquiry as the critical fundamentalist tribal discussion of humanity, as a vital part of all our lives, as *a personal and social reality, as a part of the world*, designed to help us create wiser ways of living, wiser institutions, a wiser world.

Whereas the philosophy of wisdom, in short, in subordinating intellectual inquiry to the needs of life of value, does justice to both the pragmatic and the cultural aspects of intellectual inquiry, in a

⁸⁵ Schilpp (1969, p. 80).

⁸⁶ Przibram (1967, p. 39).

unified way,⁸⁷ the philosophy of knowledge fails to do justice to both aspects.

Specialism is a relatively recent phenomenon, a general intellectual malaise that has progressively overtaken scientific, academic inquiry during the last hundred years, and especially during the last fifty years. The natural philosophers of the seventeenth century, the *philosophes* of the eighteenth century, and many scientists, philosophers, and social thinkers of the nineteenth century had no difficulty in conceiving and pursuing intellectual inquiry in broadly critical fundamentalist terms (even if epistemological and methodological misconceptions prevented them from having a full understanding of the rationale for universalism indicated here).

I have argued in this last section that there is nevertheless an even deeper intellectual and humanitarian malaise inherent in scientific, academic inquiry, which cannot by any means be construed as a relatively recent phenomenon. On the contrary, it goes back to the origins of modern science some four hundred years ago and can even be traced back to the ancient Greeks of over 2,000 years ago. It is built into the very foundations of the Western tradition. It can be put like this. Intellectual inquiry has been pursued in accordance, not with the philosophy of wisdom, but rather with the philosophy of knowledge. Instead of problems (3) and (4) of Section II being taken as fundamental, problems (1) and (2) being tackled as an aspect of, and subordinate to, problems (3) and (4), on the contrary scientific, academic inquiry has been devoted primarily to solving problems (1) and (2), solutions to aspects of these problems incidentally helping people in social life to develop improved answers to problems (3) and (4) (or so it is hoped). Instead of problems (3) and (4) being held to constitute the central problems of intellectual inquiry, on the contrary these problems have been ostracized from rational inquiry, relegated to the domain of the personal and the political, solutions to

⁸⁷ For a development of this point see my *What's Wrong With Science?* (1976a), the subtitle of which reads: *Towards a People's Rational Science of Delight and Compassion*. See too my (1984 or 2007).

them being determined by such 'irrational' factors as subjective emotion and motivation, political power, market forces.

But if present-day scientific, academic inquiry really is damagingly irrational in the quite elementary and fundamental way indicated, it may be asked: How is it possible? How can such a wholesale, fundamental irrationality have been tolerated for so long? It is not difficult to understand why in the seventeenth century questions concerning the value of life should not have been open to rational discussion: the combined power of church and state made it impossible. (One only has to remember the difficulties encountered by Galileo, Descartes, and others in seeking to establish the principle that relatively neutral problems concerning the nature of the material universe should be open to non-authoritarian, rational discussion to realize that any attempt to establish an analogous principle in connection with problems concerning the meaning and value of life was, at the time, out of the question.) The *philosophes* of the eighteenth century sought to devote reason to the enhancement of human enlightenment, human progress; unfortunately, and understandably, being over-impressed by Bacon and Newton, they failed to emphasize, clearly and unambiguously, that intellectual priority needs to be given to wisdom rather than to knowledge. Romantic writers of the late eighteenth and nineteenth centuries can be interpreted as emphasizing the priority of questions concerning life of value. Unfortunately, in doing so, they abandoned 'reason' under the mistaken impression that reason is relevant only for the acquisition of impersonal knowledge of truth, and that it involves the repression of personal feelings, desires, and imagination. The question we need to ask is this: Why have these past failures not been put right in the twentieth century? A major part of the answer is, I suggest, the increasing prevalence of specialism, which has cancelled the very possibility of critical, influential discussion of fundamentals. Indeed, the existing fundamental disorganization of contemporary scientific, academic inquiry, with its elevation of knowledge above wisdom, is just what one would expect from putting specialism into practice – as the house analogy indicates. Indeed the pursuit of knowledge dissociated from the pursuit of wisdom is itself the outcome of a kind of specialism – the

tackling of impersonal, objective, or institutional problems of knowledge dissociated from those more fundamental personal and interpersonal problems that face us in our search for what is, or can be, of value in existence. This elementary irrationality inherent in our official, public thinking about the world and ourselves is at the root of our present failure, as indicated by Higgins and others, to tackle our fundamental problems effectively and humanely. It is this that is responsible for the 'glaring contrast' noticed by Martin Luther King.

Afterword 2007

Rereading this chapter, 30 years after its initial publication in 1980, I am struck by how little has changed, how relevant what I said 30 years ago still is to academia today. "Interdisciplinarity" is, it must be admitted, now a buzz word, whereas this was hardly the case in 1980. But this seems at most to involve setting up interdisciplinary courses of various kinds, and perhaps creating interdisciplinary research groups, centres, and journals. What it does not involve is transforming the overall structure and character of universities so that an arena is created for the sustained, informal, imaginative and critical discussion of fundamental problems – discussion that influences and is influenced by more specialized research. Nor does it involve the kind of radical transformation of the whole relationship between the university and society that is required by universalism (or critical specio-fundamentalism).

In some respects, things have got worse since 1980. There has been the growth of various anti-rationalist creeds within sections of academia: post-modernism, the "strong programme" within the sociology and history of science, social constructivist views about scientific knowledge. This in turn has led some scientists and philosophers of science to defend orthodox conceptions of science and reason. It all came to a head with the publication of Alan Sokal spoof article, "Transgressing the Boundaries" (reprinted in Sokal and Bricmont (1998). The so-called "science wars" that

resulted⁸⁸ has amounted to little more than a distraction from what really does need to be debated: what kind of inquiry can best help us learn how to realize what is of value in life, for ourselves and others? What kind of inquiry can best help humanity learn how to create a better world – or at least learn how to avoid some of the worst possible future worlds?

Another regrettable development is the loss of what one might call the idea of the liberal university. When I started out as a young academic, in the mid 1960s, the idea was still around that a university should concentrate on first, hiring good people, and then, second, giving them the freedom to teach and do research as they themselves saw fit, the job of the administration being to provide support for these two essential university activities. Research assessments, committee work, and administration seem together to have all but destroyed this idea of the liberal university.

On the other hand, there have been a number of recent developments (at the time of writing) in the UK – and no doubt elsewhere – which can perhaps be interpreted as constituting first steps towards putting critical fundamentalism into academic practice. A number of new departments, institutes and centres have been created devoted to policy and peace studies. Growing concern about environmental problems, especially those associated with climate change, have led to the founding of new institutions which seek to bring specialists together to engage in relevant interdisciplinary research, and to communicate with government, the media and the public. Thus at Cambridge University there is the “Cambridge Environmental Initiative” (CEI), launched in December 2004, which brings together diverse specialized fields of research to work on environmental problems, and holds seminars and public lectures to put research scientists in touch with one another, and with the public. A similar coordinating,

⁸⁸ Provocations of or contributions to the “science wars” include: Barnes (1974); Barnes, Bloor and Henry (1996); Bloor (1976); Brown (2001); Feyerabend (1978) and (1987); Gross and Levitt (1994); Gross, Levitt and Lewis (1996); Harding (1986); Koertge (1998); Latour (1987); Pickering (1984); Segerstrale (2000); Shapin (1994); Sokal (2008); Sokal and Bricmont (1998).

interdisciplinary initiative exists at Oxford University, called the School of Geography and the Environment, founded in 2005 under another name. At University College London, my own university, there is a recent and very active initiative called the “Grand Challenges” programme, which seeks to bring together a wide range of specialists to work on four broad themes all having to do with human wellbeing. A policy document produced in November 2010 is called “Developing a Culture of Wisdom at UCL”. There is also the John Tyndall Centre for Climate Change Research, founded by 28 scientists from 10 different universities or institutions in 2000. It is based in six British universities, has links with six others, and is funded by three research councils, NERC, EPSRC and ESRC (environment, engineering and social economic research). In recent years many scientists have become concerned to involve the public in debate about questions of science policy.⁸⁹ There is now an active movement which seeks to promote public engagement with science: see note 24.

There are, in short, a few scattered signs that the revolution, from specialism to critical fundamentalism, or from knowledge to wisdom, is already under way. It will need, however, much wider cooperative support – from scientists, scholars, students, research councils, university administrators, vice chancellors, teachers, the media and the general public – if it is to become anything more than what it is at present, a few fragmentary, scattered changes intended to put right quite specific perceived defects in the *status quo*. What we need is a high profile campaign, in the public eye, concerned to make out the case for a comprehensive revolution in our universities so they come to put universalism, or the philosophy of wisdom, into academic practice. If this revolution ever comes about it will be comparable in its long-term impact to that of the Renaissance, the scientific revolution, or the Enlightenment. The outcome will be that we will at last have, what we so urgently need, institutions of learning and research rationally organized and devoted to helping us realize what is of value in life – helping us make progress towards as good a world as possible.

⁸⁹ For further details see my (2009a).

Chapter Four

Popper, Kuhn, Lakatos and Aim-Oriented Empiricism

(First published in *Philosophia*, vol. 32, 2005, pp. 181-239.)

Abstract

In this paper I argue that aim-oriented empiricism (AOE), a conception of natural science that I have defended at some length elsewhere, is a kind of synthesis of the views of Popper, Kuhn and Lakatos, but is also an improvement over the views of all three. Whereas Popper's falsificationism protects metaphysical assumptions implicitly made by science from criticism, AOE exposes all such assumptions to sustained criticism, and furthermore focuses criticism on those assumptions most likely to need revision if science is to make progress. Even though AOE is, in this way, more Popperian than Popper, it is also, in some respects, more like the views of Kuhn and Lakatos than falsificationism is. AOE is able, however, to solve problems which Kuhn's and Lakatos's views cannot solve.

1 Introduction

In this paper I argue that aim-oriented empiricism (AOE), a conception of natural science that I have spelled out and defended at some length elsewhere,⁹⁰ is a kind of synthesis of the views of Popper, Kuhn and Lakatos, but is also an improvement over the views of all three.

AOE stems from the observation that theoretical physics persistently accepts unified theories, even though endlessly many empirically more successful, but seriously disunified, *ad hoc* rivals

⁹⁰. The version of AOE defended here is a simplification and improvement of the version expounded in Maxwell (1998), in turn an improvement of versions of the view expounded in Maxwell (1972a, 1974, 1979, 1984, 1993 and 1997). For summaries of (1998) see Maxwell (1999, 2000b, 2002a, 2002b).

can always be concocted. This persistent preference for and acceptance of unified theories, even against empirical considerations, means that physics makes a persistent untestable (metaphysical) assumption about the universe: the universe is such that no seriously disunified, *ad hoc* theory is true. Intellectual rigour demands that this substantial, influential, highly problematic and implicit assumption be made explicit, as a part of theoretical scientific knowledge, so that it can be critically assessed, so that alternative versions can be considered, in the hope that this will lead to an improved version of the assumption being developed and accepted. Physics is more rigorous when this implicit assumption is made explicit even though there is no justification for holding the assumption to be true. Indeed, it is above all when there is no such justification, and the assumption is substantial, influential, highly problematic, and all too likely to be false, that it becomes especially important to implement the above requirement for rigour, and make the implicit (and probably false) assumption explicit.

Once it is conceded that physics does persistently assume that the universe is such that all seriously disunified theories are false, two fundamental problems immediately arise. What precisely ought this assumption to be interpreted to be asserting about the universe? Granted that the assumption is a pure conjecture, substantial and influential but bereft of any kind of justification, and thus all too likely in its current form to be false, how can rival versions of the assumption be rationally assessed, so that what is accepted by physics is improved?

AOE is designed to solve, or help solve, these two problems. The basic idea is that we need to see physics (and science more generally) as making not one, but a hierarchy of assumptions concerning the unity, comprehensibility and knowability of the universe, the assumptions becoming less and less substantial as one goes up the hierarchy, and thus becoming more and more likely to be true: see diagram. The idea is that in this way we separate out what is most likely to be true, and not in need of revision, at and near the top of the hierarchy, from what is most likely to be false,

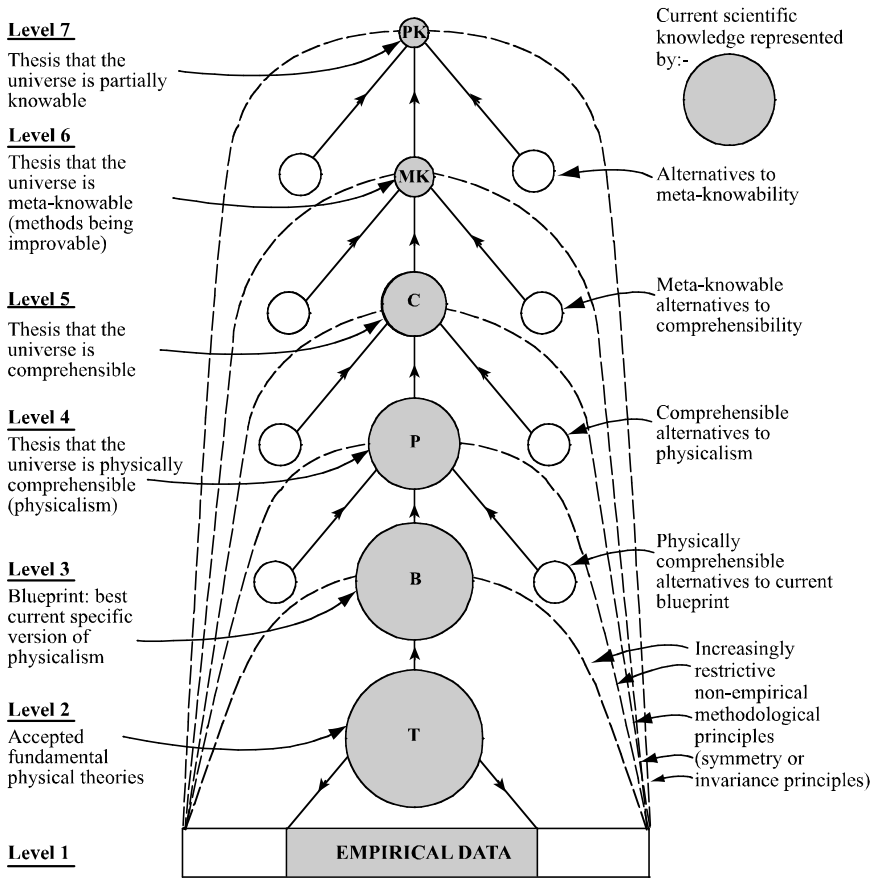


Diagram 2: Aim-Oriented Empiricism

and most in need of criticism and revision, near the bottom of the hierarchy. Evidence, at level 1, and assumptions high up in the hierarchy, are rather firmly accepted, as being most likely to be true (although still open to revision): this is then used to criticize, and to try to improve, theses at levels 2 and 3 (and perhaps 4), where falsity is most likely to be located. At the top there is the relatively insubstantial assumption that the universe is such that we can acquire some knowledge of our local circumstances. If this assumption is false, we will not be able to acquire knowledge whatever we assume. We are justified in accepting this assumption

permanently as a part of our knowledge, even though we have no grounds for holding it to be true. As we descend the hierarchy, the assumptions become increasingly substantial and thus increasingly likely to be false. At level 5 there is the rather substantial assumption that the universe is comprehensible in some way or other, the universe being such that there is just one kind of explanation for all phenomena. At level 4 there is the more specific, and thus more substantial assumption that the universe is *physically* comprehensible, it being such that there is some yet-to-be-discovered, true, unified “theory of everything”. At level 3 there is the even more specific, and thus even more substantial assumption that the universe is physically comprehensible in a more or less specific way, suggested by current accepted fundamental physical theories. Examples of assumptions made at this level, taken from the history of physics, include the following. The universe is made up of rigid corpuscles that interact by contact; it is made up of point-atoms that interact at a distance by means of rigid, spherically-symmetrical forces; it is made up of a unified field; it is made up of a unified quantum field; it is made up of quantum strings. Given the historical record of dramatically changing ideas at this level, and given the relatively highly specific and substantial character of successive assumptions made at this level, we can be reasonably confident that the best assumption available at any stage in the development of physics at this level will be false, and will need future revision. At level 2 there are the accepted fundamental theories of physics, currently general relativity and the standard model. Here, if anything, we can be even more confident that current theories are false, despite their immense empirical success. This confidence comes partly from the vast empirical content of these theories, and partly from the historical record. The greater the content of a proposition the more likely it is to be false; the fundamental theories of physics, general relativity and the standard model have such vast empirical content that this in itself almost guarantees falsity. And the historical record backs this up; Kepler’s laws of planetary motion, and Galileo’s laws of terrestrial motion are corrected by Newtonian theory, which is in turn corrected by special and general relativity; classical physics is corrected by quantum theory, in turn corrected

by relativistic quantum theory, quantum field theory and the standard model. Each new theory in physics reveals that predecessors are false. Indeed, if the level 4 assumption of AOE is correct, then all current physical theories are false, since this assumption asserts that the true physical theory of everything is unified, and the totality of current fundamental physical theory, general relativity plus the standard model, is notoriously disunified.

Finally, at level 1 there are accepted empirical data, low level, corroborated, empirical laws.

In order to be acceptable, an assumption at any level from 6 to 3 must (as far as possible) be compatible with, and a special case of, the assumption above in the hierarchy; at the same time it must be (or promise to be) empirically fruitful in the sense that successive accepted physical theories increasingly successfully accord with (or exemplify) the assumption. At level 2, those physical theories are accepted which are sufficiently (a) empirically successful and (b) in accord with the best available assumption at level 3 (or level 4). Corresponding to each assumption, at any level from 7 to 3, there is a methodological principle, represented by sloping dotted lines in the diagram, requiring that these lower down in the hierarchy are compatible with the given assumption.

When theoretical physics has completed its central task, and the true theory of everything, T, has been discovered, then T will (in principle) successfully predict all empirical phenomena at level 1, and will entail the assumption at level 3, which will in turn entail the assumption at level 4, and so on up the hierarchy. As it is, physics has not completed its task, T has not (yet) been discovered, and we are ignorant of the nature of the universe. This ignorance is reflected in clashes between theses at different levels of AOE. There are clashes between levels 1 and 2, 2 and 3, and 3 and 4. The attempt to resolve these clashes drives physics forward.

In seeking to resolve these clashes between levels, influences can go in both directions. Thus, given a clash between levels 1 and 2, this may lead to the modification, or replacement of the relevant theory at level 2; but, on the other hand, it may lead to the discovery that the relevant experimental result is not correct for any of a number of possible reasons, and needs to be modified. In

general, however, such a clash leads to the rejection of the level 2 theory rather than the level 1 experimental result; the latter are held onto more firmly than the former, in part because experimental results have vastly less empirical content than theories, in part because of our confidence in the results of observation and direct experimental manipulation (especially after expert critical examination). Again, given a clash between levels 2 and 3, this may lead to the rejection of the relevant level 2 theory (because it is disunified, *ad hoc*, at odds with the current metaphysics of physics); but, on the other hand, it may lead to the rejection of the level 3 assumption and the adoption, instead, of a new assumption (as has happened a number of times in the history of physics, as we have seen). The rejection of the current level 3 assumption is likely to take place if the level 2 theory, which clashes with it, is highly successful empirically, and furthermore has the effect of increasing unity in the totality of fundamental physical theory overall, so that clashes between levels 2 and 4 are decreased. In general, however, clashes between levels 2 and 3 are resolved by the rejection or modification of theories at level 2 rather than the assumption at level 3, in part because of the vastly greater empirical content of level 2 theories, in part because of the empirical fruitfulness of the level 3 assumption (in the sense indicated above).

It is conceivable that the clash between level 2 theories and the level 4 assumption might lead to the revision of the latter rather than the former. This happened when Galileo rejected the then current level 4 assumption of Aristotelianism, and replaced it with the idea that “the book of nature is written in the language of mathematics” (an early precursor of our current level 4 assumption). The whole idea of AOE is, however, that as we go up the hierarchy of assumptions we are increasingly unlikely to encounter error, and the need for revision. The higher up we go, the more firmly assumptions are upheld, the more resistance there is to modification.

AOE is put forward as a framework which makes explicit metaphysical assumptions implicit in the manner in which physical theories are accepted and rejected, and which, at the same time, facilitates the critical assessment and improvement of these assumptions with the improvement of knowledge, criticism being

concentrated where it is most needed, low down in the hierarchy. Within a framework of relatively insubstantial, unproblematic and permanent assumptions and methods (high up in the hierarchy), much more substantial, problematic assumptions and associated methods (low down in the hierarchy) can be revised and improved with improving theoretical knowledge. There is something like positive feedback between improving knowledge and improving (low-level) assumptions and methods - that is, knowledge-about-how-to-improve-knowledge. Science adapts its nature, its assumptions and methods, to what it discovers about the nature of the universe. This, I suggest, is the nub of scientific rationality, and the methodological key to the great success of modern science.

The above is intended to be an introductory account of AOE; further clarifications and details will emerge below when I come to expound AOE again during the course of arguing that the position can be construed to be a kind of synthesis of, and improvement over, the views of Popper, Kuhn and Lakatos.

In what follows I begin with Karl Popper, and argue that AOE can be seen to emerge as a result of modifying Popper's falsificationism⁹¹ to remove defects inherent in that position. AOE does not, however, break with the spirit of Popper's work; far from committing the Popperian sin of "justificationism", AOE is even more Popperian than Popper, in that it is a conception of science which exposes more to effective criticism than falsificationism does. Falsificationism, in comparison, shields substantial, influential and problematic scientific assumptions from criticism within science. Whereas falsificationism fails to solve what may be called the "methodological" problem of induction, AOE successfully solves the problem. And, associated with that success, AOE also solves the problem of what it means to assert of a physical theory that it is "simple", "explanatory" or "unified", a problem which falsification-ism fails to solve.

The conception of science expounded by Thomas Kuhn in his *The Structure of Scientific Revolutions* (1970a) shares important elements with Popper's falsificationism. The big difference is that whereas Kuhn holds that "normal science" is an important, healthy

2. See Popper (1959, 1963, 1983)

and entirely rational (indeed, the most rational) part of science, Popper regards normal science as "dogmatic", the result of bad education and "indoctrination", something that is "a danger to science and, indeed, to our civilization" (Popper, 1970, p. 53). It is the apparent persistent dogmatism of normal science - the persistent retention of the current paradigm in the teeth of ostensible empirical refutations - that is so irrational, so unscientific, when viewed from a falsificationist perspective. AOE, however, though subjecting scientific assumptions to even greater critical scrutiny than Popper's falsificationism, turns out to have features which are, in some respects, closer to Kuhn than to Popper. For, according to AOE, substantial and influential metaphysical assumptions are persistently accepted as a part of scientific knowledge in a way which seems much closer to the way paradigms are accepted, according to Kuhn, during normal science, than to the way falsifiable theories are to be treated in science, according to Popper. AOE depicts science as, quite properly, proceeding in a way that is reminiscent, in important respects, of Kuhn's normal science, something that is anathema to Popper's falsificationism. At the same time, AOE is free of some of the serious defects inherent in Kuhn's conception of science. Even though AOE science mimics some aspects of Kuhnian normal science, it nevertheless entirely lacks the harmful dogmatism of this kind of science, and avoids problems that arise from Kuhn's insistence that successive paradigms are "incommensurable".

Imre Lakatos's "methodology of scientific research programmes"⁹² was invented, specifically, to do justice both to Popper's insistence on the fundamental importance of subjecting scientific theories to persistent, ruthless attempted empirical refutation, and to Kuhn's insistence on the importance of preserving accepted paradigms from refutation, scientists, not paradigms, being under test when ostensible refutations arise. It is, like AOE, a kind synthesis of the ideas of Popper and Kuhn. Just as AOE incorporates elements of Popper and Kuhn, so too it incorporates elements of Lakatos's research programme methodology. At the same time, AOE is an improvement over

⁹² See Lakatos (1970, 1978).

Lakatos's view; it solves problems which Lakatos's view is unable to solve. Whereas Lakatos's view provides no means for the assessment of "hard cores" (Lakatos's "paradigms") other than by means of the empirical success and failure of the research programmes to which they give rise, AOE specifies a way in which "hard cores" (or their equivalent) can be rationally, but fallibly assessed, independent of the kind of empirical considerations to which Lakatos is restricted. This has important implications for the question of whether or not there is a rational method of discovery. It also has important implications for the *strength* of scientific method. For Lakatos, notoriously, scientific method could only decide which of two competing research programmes was the better long after the event, when one had proved to be vastly superior, empirically, to the other. "The owl of Minerva flies at dusk", as Lakatos put it, echoing Hegel. AOE provides a much more decisive methodology than Lakatos's, one which is able to deliver verdicts when they are needed, and not long after the event.

It may be thought that yet another critique of Popper, Kuhn and Lakatos is unnecessary, given the flood of literature that has appeared on the subject in the last 30 years or so: for an excellent recent survey article see Nola and Sankey (2000). My reply to this objection comes in two parts.

First, nowhere in this large body of critical literature can one find the particular line of criticism developed in the present paper. This line of criticism is, furthermore, especially fundamental and insightful in that it reveals, as other criticisms do not, what needs to be done radically to *improve* the views of Popper, Kuhn and Lakatos. Second, the improved view, namely AOE, that emerges from the criticism to be expounded here, has been entirely overlooked by the body of literature discussing and criticizing Popper, Kuhn and Lakatos. This is the decisive point. It is not enough merely to show that the views of Popper, Kuhn and Lakatos are defective. What really matters is to develop a view that overcomes these defects. That is what I set out to do here.

It is also true that, during the last 30 years, a substantial body of work has emerged on scientific method quite generally. I have in mind such publications as Holton (1973), Feyerabend (1978), Glymour (1980), van Fraassen (1980), Laudan (1984), Watkins

(1984), Hooker (1987), Hull (1988), Howson and Urbach (1993), Kitcher (1993), Musgrave (1993), Dupré (1995), McAllister (1996), Cartwright (1999). In none of these works does one find the criticism of Popper, Kuhn and Lakatos, expressed below, or the synthesis, namely AOE, which emerges from this criticism. Furthermore, the methodological views developed in the works just cited all fall to the line of criticism deployed against Popper, Kuhn and Lakatos in the present paper. There is no space to develop this last point here: it is however spelled out in Maxwell (1998, ch. 2). One implication, then, of the present paper is that philosophy of science took a wrong turning around 1974 when it failed to take up the line of argument of this paper, an early version of which is to be found in Maxwell (1974).

2 Karl Popper

As everyone knows, Popper held that science proceeds by putting forward empirically falsifiable conjectures which are then subjected to severe attempts at falsification by means of observation and experiment. Scientific theories cannot be verified by experience, but they can be falsified. Once a theory is falsified, scientists have the task of developing a potentially better theory, even more falsifiable than its predecessor, at least as ostensibly empirically successful as its predecessor, and such that it is corroborated where its predecessor was falsified. In order to be accepted (tentatively) as a part of conjectural scientific knowledge a theory must (at least) be empirically falsifiable. Non-falsifiable, metaphysical theses are meaningful, and may influence the direction of scientific research. There can even be what Popper has called "metaphysical research programmes" - programmes of research "indispensable for science, although their character is that of metaphysical or speculative physics rather than of scientific physics ... more in the nature of myths, or of dreams, than of science" (Popper, 1982a, p. 165). For Popper, metaphysical (that is, unfalsifiable) theses cannot be a part of (conjectural) scientific knowledge; such theses cannot help determine what is accepted and rejected as (conjectural) scientific knowledge, but they can influence ideas, choice of research aims and problems, in the

context of scientific discovery. For further details see Popper (1959, 1963, 1983).

Popper defended two distinct versions of falsificationism which, echoing terminology of Maxwell (1998), I shall call bare and dressed falsificationism. According to bare falsificationism, defended in Popper (1959), only empirical considerations, and such things as the falsifiability of theories and degrees of falsifiability, decide what is to be accepted and rejected in science. According to dressed falsificationism, a new theory, in order to be acceptable, "should proceed from some *simple, new, and powerful, unifying idea* about some connection or relation (such as gravitational attraction) between hitherto unconnected things (such as planets and apples) or facts (such as inertial and gravitational mass) or new "theoretical entities" (such as field and particles)" (Popper, 1963, p. 241). This "requirement of simplicity" (as Popper calls it) is in addition to anything specified in Popper (1959). In his (1959), Popper does, it is true, demand of a theory that it should be as simple as possible, but Popper there identifies degree of simplicity of a theory with degree of falsifiability. (There is a second, related notion, but Popper makes it clear that if the two clash it is the falsifiability notion, just indicated, which takes priority: see page 130). Thus, in his (1959), in requiring of an acceptable theory that it should be as simple as possible, Popper is demanding no more than that it should be as falsifiable as possible. But Popper's "requirement of simplicity" of his (1963) is wholly in addition to falsifiability. A theory of high falsifiability may not "proceed from some simple, new, and powerful unifying idea", and vice versa. We thus have two versions of falsificationism before us: bare falsificationism of Popper's (1959), and dressed falsificationism of (1963, chapter 10), with the new "requirement of simplicity" added onto the (1959) doctrine.

I now give my argument for holding that neither doctrine is tenable. My argument is *not* that Popper fails to show how theories can be verified, or rendered probable; nor is my argument that Popper fails to show how scientific theories can be falsified, in that falsification requires the *verification* of a low-level falsifying

hypothesis (which, according to Popper, is not possible).⁹³ There is nothing "justificationist", in other words, about my criticism. It amounts simply to this. Bare falsificationism fails dramatically to do justice to the way theories are selected in science (entirely independently of any question of verification, justification or falsification). Dressed falsificationism does better justice to scientific practice, but commits science to making substantial, influential and problematic assumptions that remain implicit, and cannot adequately be made explicit within science. Science pursued in accordance with dressed falsificationism is irrational, in other words, because it fails to implement the elementary, and quasi-Popperian, requirement for rationality that "assumptions that are substantial, influential, problematic and implicit need to be made explicit, so that they can be critically assessed and so that alternatives may be put forward and considered, in the hope that such assumptions can be improved" (Maxwell, 1998, p. 21). Dressed falsificationism fails, in other words, for good Popperian reasons: it fails to expose substantial, influential, problematic assumptions to *criticism* within science.

3 Refutation of Bare Falsificationism

Here, then, in a little more detail, is my refutation of bare falsificationism. Given any accepted physical theory, at any stage in the development of physics, however empirically successful (however highly corroborated) - Newtonian theory, say, or classical electrodynamics, quantum theory, general relativity, quantum electrodynamics, chromodynamics or the standard model - there will always be endlessly many rival falsifiable theories that can easily be formulated which will fit the available data just as well as the accepted theory. Taking Newtonian theory (NT) as an example of an accepted theory, here are two examples of rival theories. NT*: "Everything occurs as NT asserts, until the first second of 2100, when an inverse cube law of gravitation will abruptly hold". NT**: "Everything occurs as NT asserts, except for systems consisting of gold spheres, each having a mass of 1,000

4 For Popper's replies to such criticisms: see Popper (1972), chapter 1; (1974), sections II and III; and (1983), Introduction and chapter 1.

tons, interacting with each other gravitationally in outer space, in a vacuum, within a spherical region of 10 miles: for these systems, Newton's law of gravitation is repulsive, not attractive". (For further examples and discussion, see Maxwell, 1998, pp. 47-54). It is easy to see that there are infinitely many such rivals to NT, just as empirically successful (at the moment) as NT. The predictions of NT may be represented as points in a multi-dimensional space, each point corresponding to some specific kind of system (there being infinitely many points). NT has only been verified (corroborated) for a minute region of this space. In order to concoct a (grossly *ad hoc*) rival to NT, just as empirically successful as NT, all we need do is identify some region in this space that includes no prediction of NT that has been verified, and then modify the laws of NT arbitrarily, for just that identified region.

The crucial question now is this: on what basis does bare falsificationism reject all these falsifiable but unfalsified rival theories? According to bare falsificationism, T_2 is to be accepted in preference to T_1 if T_1 has been falsified, T_2 has greater empirical content (is more falsifiable) than T_1 , T_2 successfully predicts all that T_1 successfully predicts, T_2 successfully predicts the phenomena that falsified T_1 , and T_2 successfully predicts new phenomena not predicted by T_1 (see Popper, 1959, pp. 81-84 and elsewhere). Given NT, it is a simple matter to concoct rival theories, of the above type, that satisfy all the above bare falsificationist requirements for being more acceptable than NT. Most accepted physical theories yield empirical predictions that clash with experiments, and thus are ostensibly falsified. We can always concoct new theories, in the way just indicated, doctored to yield the "correct" predictions. We can add on independently testable auxiliary postulates, thus ensuring that the new theory has greater empirical content than the old one. And no doubt this excess content will be corroborated. For details of how this can be done see Maxwell (1998, pp. 52-54). Such theories are, of course, grossly *ad hoc*, grossly "aberrant" as I have called them; but they satisfy Popper's (1959) requirements for being better theories than accepted physical theories.

It is worth noting that such "better" theories need not be quite as wildly *ad hoc* as the ones indicated above; sometimes such theories are actually put forward in the scientific literature, and yet are not taken seriously, even by their authors, let alone by the rest of the scientific community. An example is an *ad hoc* version of NT put forward by Maurice Levy in 1890, which combined in an *ad hoc* way two distinct modifications of Newton's law of gravitation, one based on the way Weber had proposed Coulomb's law should be modified, the other based on the way Riemann had proposed Coulomb's law should be modified: for details see North (1965). By 1890, NT had been refuted by observation of the precession of the perihelion of the orbit of Mercury; attempts to salvage NT by postulating an additional planet, Vulcan, had failed. Levy's theory successfully predicted all the success of NT, and in addition successfully predicted the observed orbit of Mercury, just that which refuted NT; in addition, of course, it made predictions different from NT for further Sun-Mercury type systems not yet observed. Despite this, Levy's theory was not taken seriously for a moment, not even by Levy himself. How can bare falsificationism recommend rejection of such *ad hoc* versions of NT when they satisfy all the requirements of bare falsificationism for being more acceptable theories? No adequate answer is forthcoming, and it is this which spells the downfall of bare falsificationism (as Popper may himself have realized when he put forward dressed falsificationism in his (1963), chapter 10).

Note, again, that this criticism of Popper has nothing justificational about it whatsoever: it simply points to the drastic failure of bare falsificationism to do justice to what actually goes on in physics.

It may be objected that *ad hoc* rivals to NT of the kind just considered are so silly, so crackpot, that they do not *deserve* to be taken seriously within physics.⁹⁴ This is of course correct. The

⁹⁴ Popper discusses such "silly" rival theories in Popper (1983, pp. 67-71). He argues that they deserve to be rejected on the grounds that they create more problems than they solve, problems of explanation. This is a relevant consideration granted dressed falsificationism, but not granted bare falsificationism. He also argues that it does not matter if such "silly" theories become potential rivals, since it can be left to scientists themselves to criticize

crucial point, however, is that bare falsificationism ought to be able to deliver this verdict, and this it singularly fails to do. Bare falsificationism actually declares of appropriately concocted *ad hoc* rivals to NT that these are *better, more acceptable* than NT.

But can a criticism of Popper that appeals to such silly, crackpot theories be taken seriously? I have two replies to this question. First, not all the *ad hoc* or aberrant variants are entirely silly. Levy's theory is perhaps an example. There are degrees of *ad hocness*, from the utterly crackpot and absurd, to a degree of *ad hocness*, so slight, so questionable, in comparison, that the issue of whether the theory really is *ad hoc* or not may be hotly disputed by physicists themselves. (Such disputes arise especially during scientific revolutions.) This is an important point which will have a bearing on the argument of the next section. Second, it is, I submit, the very silliness of these crackpot theories that makes the above criticism of Popper so serious. If bare falsificationism favoured T_1 over T_2 , while most scientists favoured T_2 over T_1 , even though admitting that T_1 is nevertheless a good theory, almost as acceptable as T_2 , bare falsificationism would not be in such trouble. What is lethal for bare falsificationism is that it declares T_1 to be better than T_2 in circumstances where scientists themselves (and all of us) can see that T_2 is vastly superior to T_1 , T_1 being grossly *ad hoc*, aberrant, wholly crackpot and silly. Bare falsificationism favours theories that receive, and deserve, instant

them. But what this ignores is that it is precisely Popper's methodology which should be providing guidelines for such criticism. Far from condemning such a "silly" theory as worthy of rejection, bare falsificationism holds such a theory to be better than the accepted theory (if it has greater empirical content, is not falsified where the accepted theory appears to be, and some of the excess content of the "silly" theory is corroborated). Popper fails to appreciate that it his methodology, not he himself, which needs to declare that silly theories are indeed "silly". The fact that his methodology declares these silly theories to be highly acceptable is a devastating indictment of his methodology. To argue that these silly theories, refuting instances of his methodology, do not matter and can be discounted, is all too close to a scientist arguing that evidence, that refutes his theory, should be discounted, something which Popper resoundingly condemns. The falsificationist stricture that scientists should not discount falsifying instances, ought to apply to methodologists as well!

rejection: there could scarcely be a more decisive falsification of falsificationism than that.

4 Refutation of Dressed Falsificationism

Having argued that Popper's (1959) bare falsificationism is untenable, I turn my attention now to Popper's (1963, chapter 10) doctrine of dressed falsificationism. As I have mentioned, this adds onto the (1959) doctrine Popper's new "requirement of simplicity" (Popper, 1963, p. 241): see section 2 above.

As long as there is no serious ambiguity as to what proceeding "from some simple, new, and powerful, unifying idea" means, it is at once clear that the new doctrine is able to exclude from science all the empirically successful but *ad hoc*, aberrant, crackpot, silly theories, of the kind discussed above. They do not proceed "from some simple...unifying idea", and are to be rejected on that account, whatever their empirical success may be, even if this empirical success is greater than accepted scientific theories.

However, adopting Popper's new "principle of simplicity" as a basic methodological principle of science has the effect of permanently excluding from science all *ad hoc* theories that fail to satisfy the principle, however empirically successful such theories might be if considered. This amounts to assuming permanently that the universe is such that no *ad hoc* theory, that fails to satisfy Popper's principle of simplicity, is true. It amounts to accepting, as a permanent item of scientific knowledge, the substantial metaphysical thesis that the universe is non-*ad hoc*, in the sense that no theory that fails to satisfy Popper's principle of simplicity is true, however empirically successful it might turn out to be if considered. But this, of course, clashes with Popper's criterion of demarcation: that no unfalsifiable, metaphysical thesis is to be accepted as a part of scientific knowledge. If the demarcation principle is upheld, then the metaphysical thesis just indicated, asserting that the universe is non-*ad hoc*, remains implicit in the permanent adoption of Popper's principle of simplicity as a basic methodological principle of science. (And this is the way Popper himself seems to have conceived the matter: he says of metaphysical research programmes that they are "often held unconsciously", and "are implicit in the theories and in the

attitudes and judgements of the scientists": (Popper, 1982a, p. 161).) But in leaving the metaphysical thesis of non-*ad hocness* implicit in the methodological principle of simplicity, science violates an elementary requirement for rationality, already mentioned, according to which "assumptions that are substantial, influential, problematic and implicit need to be made explicit, so that they can be critically assessed and so that alternatives may be put forward and considered, in the hope that such assumptions can be improved" (Maxwell, 1998, p. 21). The non-*ad hoc* metaphysical assumption may, after all, be false. We may need to adopt a modified version of the assumption. It may be essential for the progress of science that this assumption is modified. Just this turns out to be the case, given certain formulations of the assumption, as we shall see below. In leaving the non-*ad hoc* metaphysical assumption implicit in the adoption of the methodological principle of simplicity, dressed falsificationism protects this substantial, influential and highly problematic assumption from *criticism*, from the active consideration of alternatives.⁹⁵

Dressed falsificationism fails, in other words, for good Popperian reasons: it is either inconsistent (in that the untestable, metaphysical thesis that the universe is non-*ad hoc* is held to be a part of conjectural scientific knowledge, in conflict with the principle of demarcation), or it irrationally protects an implicit,

⁹⁵. In fact even the methodological rules of bare falsificationism are such that persistent application of these rules commits one to making implicit metaphysical assumptions (which may be false). Bare falsificationism, as formulated by Popper, requires of an acceptable theory that it is strictly universal in that it makes no reference to any specific time, place or object. This makes it impossible for science to discover that the laws of nature just are different within specific space-time regions, or that there is a specific object with unique dynamical properties. There is no scope, within bare falsificationism, for the rejection of these metaphysical theses, even though circumstances could conceivably arise such that progress in knowledge would require this. (AOE, by contrast, allows for this remote possibility: that which is dogmatically upheld by bare falsificationism becomes criticizable granted AOE.) Popper recognizes that the methodological rule requiring any theory to be strictly universal does have a metaphysical counterpart (1959, sections 11 and 79), but fails to appreciate how damaging this is for falsificationism.

substantial assumption from explicit criticism within the intellectual domain of science.

Here again, it should be noted, there is nothing justificationist about this criticism of Popper's dressed falsificationism. On the contrary, what the argument shows is that dressed falsificationism protects a substantial, influential, problematic but implicit assumption from criticism within science: Popper's doctrine fails for the good Popperian reason of restricting criticism.

It may be objected that adopting Popper's methodological principle of simplicity does *not* commit science to making a substantial metaphysical assumption about the universe - namely, that it is such that no falsifiable theory, however empirically successful, which fails to satisfy the principle, is true. But I do not see how such an objection can be valid. Suppose, instead of adopting Popper's principle, science adopted the principle: in order to be acceptable, a new physical theory must postulate that the universe is made up of atoms. This methodological principle is upheld in such a way that even though theories are available which postulate fields rather than atoms, and which are much more empirically successful than any atomic theory, nevertheless these rival field theories are all excluded from science. Would it not be clear that science, in adopting and implementing the methodological principle of atomicity in this way, is making the assumption that the universe is made up of atoms, whether this is acknowledged or not? How can this be denied? Just the same holds if science adopts and implements Popper's methodological principle of simplicity.

Popper might have tried to wriggle out of accepting this conclusion by pointing to the fact that he only declared that a new theory, in order to be acceptable, "should" proceed from some simple, unifying idea. It is desirable, but not essential, that new theories should satisfy this principle. The principle is relevant to the context of discovery, perhaps, but not to the context of acceptance and rejection. (It is a heuristic principle, not a methodological one.) But if Popper's doctrine is interpreted in this way, it immediately fails to overcome the objections spelled out in section 3 above. Either falsificationism adopts Popper's principle of simplicity as a *methodological* principle, or it does not. If it

does, it encounters the objections just indicated; if it does not, it encounters the objections of section 3.

5 From Falsificationism to Aim-Oriented Empiricism

The conclusion to be drawn from the argument so far is that science is more rational, more intellectually rigorous, if it makes explicit, as a criticizable tenet of (conjectural) scientific knowledge, that substantial, influential and problematic metaphysical thesis which is implicit in the way physics persistently rejects *ad hoc* theories, however empirically successful they may be. At once two important new problems leap to our attention. What, precisely, does this metaphysical thesis assert? And on what grounds is it to be (conjecturally) accepted as a part of scientific knowledge? The conception of science which I uphold as a radical improvement over Popper's falsificationism, namely aim-oriented empiricism (AOE), is put forward as the solution to these two problems. I now expound AOE (in a little more detail than the introductory exposition of section 1) and indicate how it solves the two problems just mentioned; I indicate further how it solves the methodological problem of induction and the related problem of simplicity, and then consider possible objections.

As far as the first of the above two problems is concerned, a wide range of metaphysical theses are available. As I indicated in section 3 above, *ad hoc* theories range from the utterly crackpot and silly, to theories that are only somewhat lacking in simplicity or unity. At one extreme, we might adopt a metaphysical thesis that excludes only utterly silly theories; at the other extreme, we might adopt the thesis that the universe is physically comprehensible in the sense that it has a unified dynamic structure, some yet-to-be-discovered unified physical "theory of everything" being true - a thesis that I shall call "physicalism". We might even adopt some specific version of physicalism, which asserts that the underlying physical unity is of a specific type: it is made up of a unified field perhaps, or a quantum field, or empty topologically complex curved space-time, or a quantum string field. Other things being equal, the more specific the thesis (and thus the more it excludes) so the more likely it is to be false, whereas the more unspecific it is so the more likely it is to be true.

As far as the second of the above two problems is concerned, there are four considerations that we can appeal to, three wholly Popperian in spirit if not in the letter of Popperian doctrine.

(1) If some metaphysical thesis, M , is implicit in some scientific methodological practice, then science is more rigorous if M is made explicit, since this facilitates criticism of it, the consideration of alternatives.

(2) A metaphysical thesis may be such that its truth is a necessary condition for it to be possible for us to acquire knowledge: if so, accepting the thesis can only help, and cannot undermine, the pursuit of knowledge of truth.

(3) Given two rival metaphysical theses, M_1 and M_2 , it may be the case that M_1 supports an empirical scientific research programme that has apparently met with far greater empirical success than any rival empirical research programme based on M_2 : in this case we may favour M_1 over M_2 , at least until M_2 , or some third thesis, M_3 , shows signs of supporting an even more empirically progressive research programme.

(4) M_1 may be preferred to M_2 on the grounds that it gives greater promise of supporting an empirically progressive research programme.

The arguments of sections 3 and 4 have established that physics must accept (conjecturally) some kind of metaphysical thesis of non-*ad hocness*, if crackpot theories are to be excluded: it makes sense to adopt that thesis which seems to be the most fruitful in promoting scientific progress. (To say that M_1 "supports" an empirically successful research programme is to say that the programme develops a succession of theories, each empirically more successful than its predecessors, in a Popperian sense, and each being closer to exemplifying, to being a precise, testable instantiation of, M_1 than its predecessors.)

Two difficulties arise, however, when one attempts to use (2) and (3) to select the best available metaphysical thesis from the infinitely many options available. First, as far as (2) is concerned, any thesis sufficiently substantial to exclude empirically successful crackpot theories from science is such that acquisition of knowledge might still be possible even if the thesis is false. On the other hand, any thesis such that its truth is necessary for knowledge

to be acquired is much too insubstantial to exclude crackpot theories. Second, as far as (3) is concerned, given any metaphysical thesis, M, that supports a non-crackpot empirically progressive scientific research programme, we can mimic this with a crackpot M* that supports a crackpot empirically progressive research programme, with a series of crackpot theories, T1*, T2*, ..., these theories becoming progressively more and more empirically successful, and closer and closer to exemplifying M*.

These two difficulties can be overcome, however, if physics is construed as adopting a hierarchy of metaphysical conjectures concerning the comprehensibility and knowability of the universe, these conjectures becoming more and more insubstantial as one ascends the hierarchy, more and more likely to be true: see diagram. At level 7 there is the thesis that the universe is such that we can continue to acquire knowledge of our local circumstances, sufficient to make life possible. At level 6 there is the more substantial thesis that there is some rationally discoverable thesis about the nature of the universe which, if accepted, makes it possible progressively to improve methods for the improvement of knowledge. "Rationally discoverable", here, means at least that the thesis is not an arbitrary choice from infinitely many analogous theses. At level 5 we have the even more substantial thesis that the universe is *comprehensible* in some way or other, whether physically or in some other way. This thesis asserts that the universe is such that there is *something* (God, tribe of gods, cosmic goal, physical entity, cosmic programme or whatever), which exists everywhere in an unchanging form and which, in some sense, determines or is responsible for everything that changes (all change and diversity in the world in principle being explicable and understandable in terms of the underlying unchanging *something*). A universe of this type deserves to be called "comprehensible" because it is such that everything that occurs, all change and diversity, can in principle be explained and understood as being the outcome of the operations of the one underlying *something*, present throughout all phenomena. At level 4 we have the still more substantial thesis that the universe is *physically* comprehensible in

some way or other (a thesis I shall call *physicalism*⁹⁶). This asserts that the universe is made up one unified self-interacting physical entity (or one kind of entity), all change and diversity being in principle explicable in terms of this entity. What this amounts to is that the universe is such that some yet-to-be-discovered unified physical theory of everything is true. At level 3, we have an even more substantial thesis, the best, currently available specific idea as to how the universe is physically comprehensible. This asserts that everything is made of some specific kind of physical entity: corpuscle, point-particle, classical field, quantum field, convoluted space-time, string, or whatever. Because the thesis at this level is so specific, it is almost bound to be false (even if the universe is physically comprehensible in some way or other). Here, ideas evolve with evolving knowledge. At level 2, we have our best fundamental physical theories, currently general relativity and the so-called standard model, and at level 1 we have empirical data (low level experimental laws).

The thesis at the top of the hierarchy, at level 7, is such that, if it is false, knowledge cannot be acquired whatever is assumed. This thesis is, quite properly, accepted as a permanent part of scientific knowledge, in accordance with (2) above, since accepting it can only help, and cannot hinder, the acquisition of knowledge whatever the universe is like.

I have two arguments (appealing to (4) above) for the acceptance of the thesis of meta-knowability, at level 6.

(i) Granted that there is *some* kind of general feature of the universe which makes it possible to acquire knowledge of our local environment (as guaranteed by the thesis at level 7), it is reasonable to suppose that we do not know all that there is to be known about what the *nature* of this general feature is. It is reasonable to suppose, in other words, that we can improve our knowledge about the nature of this general feature, thus improving

⁹⁶. Smart (1963) has used the term 'physicalism' to stand for the view that the world is made up entirely of physical entities of the kind postulated by fundamental physical theories - electrons, quarks and so on. As I am using the term, 'physicalism' stands for the very much stronger doctrine that the universe is physically comprehensible, that it is such that some yet-to-be-discovered, unified "theory of everything" is true.

methods for the improvement of knowledge. Not to suppose this is to assume, arrogantly, that we already know all that there is to be known about how to acquire new knowledge. Granted that learning is possible (as guaranteed by the level 7 thesis), it is reasonable to suppose that, as we learn more about the world, we will learn more about how to learn. Granted the level 7 thesis, in other words, meta-knowability is a reasonable conjecture.

(ii) Meta-knowability is too good a possibility, from the standpoint of the growth of knowledge, not to be accepted initially, the idea only being reluctantly abandoned if all attempts at improving methods for the improvement of knowledge fail.

These two arguments for accepting meta-knowability are, admittedly, weak. It is crucial, however, that these two arguments make no appeal to the success of science, for a reason that will become apparent in a moment.

The thesis that the universe is comprehensible, at level 5 is accepted because no rival thesis, at that level, has been so fruitful in leading to empirically progressive research programmes. It is hardly an exaggeration to say that all empirically successful research programmes into natural phenomena have been organized around the search for explanatory theories, of one kind or another. Aberrant rivals to the thesis of comprehensibility, which might be construed as supporting aberrant empirically successful research programmes, are rejected because of incompatibility with the thesis of meta-knowability at level 6. Such rival ideas are not “rationally discoverable” in that each constitutes an arbitrary choice from infinitely many equivalent rivals.

Physicalism at level 4 is accepted because it is by far the most empirically fruitful thesis at that level that is compatible with the thesis of comprehensibility, at level 5.

Since the scientific revolution of the 17th century, all new fundamental physical theories have enhanced overall unity of theoretical physics. Thus Newtonian theory (NT) unifies Galileo's laws of terrestrial motion and Kepler's laws of planetary motion (and much else besides). Maxwellian classical electrodynamics, (CEM), unifies electricity, magnetism and light (plus radio, infra red, ultra violet, X and gamma rays). Special relativity (SR) brings greater unity to CEM (in revealing that the way one divides up the electromagnetic field into the electric and magnetic fields depends on one's reference frame). SR is also a step towards unifying NT and CEM in that it transforms space and time so

as to make CEM satisfy a basic principle fundamental to NT, namely the (restricted) principle of relativity. SR also brings about a unification of matter and energy, via the most famous equation of modern physics, $E = mc^2$, and partially unifies space and time into Minkowskian space-time. General relativity (GR) unifies space-time and gravitation, in that, according to GR, gravitation is no more than an effect of the curvature of space-time. Quantum theory (QM) and atomic theory unify a mass of phenomena having to do with the structure and properties of matter, and the way matter interacts with light. Quantum electrodynamics unifies QM, CEM and SR. Quantum electroweak theory unifies (partially) electromagnetism and the weak force. Quantum chromodynamics brings unity to hadron physics (via quarks) and brings unity to the eight kinds of gluon of the strong force. The standard model unifies to a considerable extent all known phenomena associated with fundamental particles and the forces between them (apart from gravitation). The theory unifies to some extent its two component quantum field theories in that both are locally gauge invariant (the symmetry group being $U(1) \times SU(2) \times SU(3)$). String theory, or M-theory, holds out the hope of unifying all phenomena. All these theories have been accepted because they progressively (a) *increase* the overall unity of theoretical physics and (b) *increase* the predictive power of physical theory, (a) being as important as (b). Physicalism is the key, persisting thesis of the entire research programme of theoretical physics since Galileo, and no obvious rival thesis, at that level of generality, can be substituted for physicalism in this research programme.

It may be asked: But how can this succession of theories reinforce physicalism when the totality of physical theory has always, up till now, clashed with physicalism? The answer: If physicalism is true, then all physical theories that only unify a restricted range of phenomena, must be false. Granted the truth of physicalism, and granted that theoretical physics advances by putting forward theories of limited but ever increasing empirical scope, then it follows that physics will advance from one false theory to another (as it has done: see point 7 of section 6 below), all theories being false until a unified theory of everything is achieved (which just might be true). The successful pursuit of physicalism requires progressive *increase* in both empirical scope and unity of the totality of fundamental physical theory. It is just this which the history of physics, from Galileo to today, exemplifies - thus demonstrating the unique fruitfulness of physicalism.

At level 3 that metaphysical thesis is accepted which is the best specific version of physicalism available, that seems to do the best justice to the evolution of physical theory. Two considerations govern acceptance of testable fundamental dynamical physical theories. Such a theory must be such that (i) it, together with all other accepted fundamental physical theories, exemplifies, or is a special case of, the best available metaphysical blueprint (at level 3), and physicalism (at level 4) to a sufficiently good extent, and

(ii) it is sufficiently successful empirically (where empirical success is to be understood, roughly, in a Popperian sense).

How does this hierarchical view of AOE overcome the problems and difficulties, indicated above, that confront any view which holds that science makes just one, possibly composite metaphysical assumption, at just one level? Given the one-thesis view, it must remain entirely uncertain as to what the one thesis should be. If it is relatively contentful and precise, more or less equivalent to the current level 3 thesis of AOE, then it is all too likely that this is false, and will need to be replaced in the future. If it is relatively contentless and imprecise, equivalent to theses at levels 7 or 6, this will not be sufficiently precise to exclude empirically successful but grossly *ad hoc*, aberrant theories. Even the level 4 thesis of physicalism is both too contentful and precise, and not contentful and precise enough. Physicalism may be false, and may need to be revised. At the same time, physicalism lacks the potential heuristic power to suggest good new fundamental theories which the more precise and contentful theses at level 3 possess. All these difficulties are avoided by the hierarchical view of AOE, just because of the hierarchy of assumptions, graded from the relatively contentless, imprecise and permanent at the top, to the relatively contentful, precise and impermanent (but methodologically and heuristically fruitful) at the bottom.

Any one-thesis view faces the even more serious problem of how this one thesis is to be critically assessed, revised, and improved. The hierarchical view of AOE overcomes this problem by providing severe constraints on *what* is to be revised, and *how* this revision is to proceed. In the first instance, and only in quite exceptional circumstances, only the current level 3 thesis can be revised. This revision must proceed, however, within constraints provided by the level 4 thesis of physicalism, on the one hand, and accepted, empirically successful level 2 theories, on the other hand. In a really exceptional situation, scientific progress might require the revision of the level 4 thesis of physicalism, but this too would proceed within the constraints of the thesis at level 5, and empirically successful theories at level 2, or empirically progressive research programmes at levels 2 and 3. The great merit of AOE is that it separates out what is most likely to be true from

what is most likely to be false in the metaphysical assumptions of physics, and employs the former to assess critically, and to constrain, theses that fall into the latter category. It concentrates criticism and innovation where it is most likely to promote scientific progress.

Finally, any one-thesis view cannot, as we have seen, simultaneously call upon principles (1) to (4) to justify acceptance of the single thesis, whatever it may be. The hierarchical view of AOE is able to do just that. It can appeal to *different* principles, (1) to (4) above, to justify⁹⁷ (to provide a rationale for) acceptance of the *different* theses at the different levels of the hierarchy of AOE. Thus acceptance of the thesis at level 7 is justified by an appeal to (2); acceptance of theses at levels 3 to 5 are accepted as a result of (a) an appeal to (3), and (b) compatibility with the thesis above in the hierarchy. The thesis at level 6 is accepted as a result of an appeal to (4). Aberrant rivals to theses accepted at levels 3 to 5 (which might be construed to support aberrant, rival empirically progressive research programmes) are excluded on the grounds that these clash with the thesis at level (6). For further details of how AOE overcomes the two difficulties indicated above, and for further details of the view itself, see Maxwell (1998, chapter 5, and elsewhere).

It may be objected that AOE suffers from vicious circularity, in that acceptance of physical theories is justified by (in part) an appeal to physicalism, the acceptance of which is justified, in turn, by the empirical success of physical theory. My reply to this objection is that the level 6 thesis of meta-knowability asserts that the universe is such that this kind of circular methodology, there being positive feedback between metaphysics, methods, and empirically successful theories, is just what we need to employ in

8. This talk of “justifying” may seem thoroughly unPopperian in character, but it is not. What is at issue is not the justification of the truth, or probable truth, of some thesis, but only the justification of *accepting* the thesis (granted our aim is truth). Within Popper’s falsificationism, there is just such a “justification” for accepting highly falsifiable (and unfalsified) theories: such theories, being most vulnerable to falsification, facilitate the discovery of error, and thus give the most hope of progress (towards truth). Acceptance of such theories is justified (according to falsificationism) because it promotes error detection and progress.

order to improve our knowledge. The thesis of meta-knowability, if true, justifies implementation of AOE. This response is only valid, of course, if reasons for accepting the level 6 thesis of meta-knowability do not themselves appeal to the success of science (which would just reintroduce vicious circularity at a higher level). As I made clear above, the two arguments given for accepting meta-knowability make no appeal to the success of science whatsoever.⁹⁸

A basic idea of AOE is to channel or direct criticism so that it is as fruitful as possible, from the standpoint of aiding progress in knowledge. The function of criticism within science is to promote scientific progress. When criticism demonstrably cannot help promote scientific progress, it becomes irrational (the idea behind (2) above). In an attempt to make criticism as fruitful as possible, we need to try to direct it at targets which are the most fruitful, the most productive, to criticize (from the standpoint of the growth of knowledge). This is the basic idea behind the hierarchy of AOE. Conjectures at all levels remain open to criticism. But, as we ascend the hierarchy, conjectures are less and less likely to be false; it is less and less likely that criticism, here, will help promote scientific knowledge. The best currently available level 3 conjecture is almost bound to be false: the history of physics reveals, at this level, as I have indicated above, that a number of different conjectures have been adopted and rejected in turn. Here, criticism, the activity of developing alternatives (compatible with physicalism) is likely to be immensely fruitful for progress in theoretical physics. Indeed, in Maxwell (1998, pp. 78-89, 159-163 and especially 217-223), I argue that this provides physics with a rational, though fallible and non-mechanical method for the discovery of new fundamental physical theories, a method invented and exploited by Einstein in discovering special and general relativity (Maxwell, 1993, pp. 275-305), something which Popper has argued is not possible: see Popper (1959, pp. 31-32). Criticizing physicalism, at level 4, may also be fruitful for physics, but (the conjecture of AOE is) that this is not as likely to be as

9. For a more detailed rebuttal of this objection see Maxwell (2004, pp. 207-210).

fruitful as criticism at level 3. (Elsewhere I have suggested an alternative to physicalism: see Maxwell, 2004, pp. 198-205.) And, as we ascend the hierarchy (so AOE conjectures), criticism becomes progressively less and less likely to be fruitful. Against that, it must be admitted that the higher in the hierarchy we need to modify our ideas, so the more dramatic the intellectual revolution that this would bring about. If physicalism is rejected altogether, and some quite different version of the level 5 conjecture of comprehensibility is adopted instead, the whole character of natural science would change dramatically; physics, as we know it, might even cease to exist.

The biggest change, in moving from falsificationism to AOE, has to do with the role of metaphysics in science, and the scope of scientific knowledge. According to falsificationism, untestable metaphysical theses may influence scientific research in the context of discovery, and may even lead to metaphysical research programmes; they cannot, however, be a part of scientific knowledge itself. But according to AOE, the metaphysical theses at levels 3 to 7 are all a part of current (conjectural) scientific knowledge. In particular, physicalism is. According to AOE, it is a part of current scientific knowledge that the universe is physically comprehensible - certainly not the case granted falsificationism.

Another important change has to do with the relationship between science and the philosophy of science. Falsificationism places the study of scientific method, the philosophy of science, outside science itself, in accordance with Popper's demarcation principle. AOE, by contrast, makes scientific method and the philosophy of science an integral part of science itself. The activity of tackling problems inherent in the aims of science, at a variety of levels, and of developing new possible aims and methods, new possible more specific or less specific philosophies of science (views about what the aims and methods of science ought to be) is, according to AOE, a vital research activity of science itself. But this is also philosophy of science, being carried on within the framework of AOE.⁹⁹

⁹⁹ In holding that metaphysical theses and philosophies of science are an integral part of science itself, AOE implies that Popper's principle of demarcation

AOE differs in many other important ways from Popper's falsificationism, whether bare or dressed (see Maxwell, 1998). Nevertheless the impulse, the intellectual aspirations and values, behind the hierarchical view of AOE are, as I have tried to indicate, thoroughly Popperian in character and spirit. The whole idea is to turn implicit assumptions into explicit conjectures in such a way that criticism may be directed at what most needs to be criticized from the standpoint of aiding progress in knowledge, so that conjectures may be developed and adopted that are the most fruitful in promoting scientific progress, at the same time no substantial conjecture, implicit or explicit, being held immune from critical scrutiny.

6 Aim-Oriented Empiricism an Improvement over Falsificationism

AOE is also, in a number of ways, a considerable improvement over Popper's falsificationism.

1. *Consistency.* Bare falsificationism fails dramatically to do justice to scientific practice, and is an inherently unworkable methodology, in any case. (In what follows I shall mostly ignore bare falsificationism as obviously untenable, and concentrate on comparing dressed falsificationism and AOE.) Dressed falsificationism does better justice to scientific practice, but at the cost of consistency; persistent rejection of empirically successful theories that do not "proceed from some simple...unifying idea" commits science to accepting a metaphysical thesis of simplicity as a part of scientific knowledge (though this is not recognized); this

(Popper, 1963, chapter 11) is to be rejected. Popper's demarcation proposal, apart from being untenable, is in any case too simplistic, in that it reduces to one a number of distinct demarcation issues. Popper rolls into one the distinct tasks of demarcating (a) good from bad science, (b) science from non-science, (c) science from pseudo-science, (d) rational from irrational inquiry, (e) knowledge from mere speculation, (f) knowledge from dogma (or superstition, or prejudice, or popular belief), (g) the empirical from the metaphysical, and (h) factual truth from non-factual (analytic) truth. (a) to (d) involve demarcating between disciplines, whereas (e) to (h) involve demarcating between propositions.

contradicts Popper's demarcation principle. AOE is free of such lethal defects.

2. *Criticism.* Pursuing physics in accordance with dressed falsificationism protects the implicit metaphysical thesis of simplicity from criticism within science itself, just because this thesis is metaphysical (and therefore not a part of science) and implicit (and therefore not available for sustained, explicit critical scrutiny). AOE, by contrast, is specifically designed to provide a framework of metaphysical assumptions and corresponding methodological rules within which level 3 metaphysical blueprints may be developed, and critically assessed, within science.

3. *Rigour.* Science pursued in accordance with AOE is more rigorous than science pursued in accordance with falsificationism. An elementary, but important requirement for rigour is that assumptions that are substantial, influential, problematic and implicit need to be made explicit so that they can be criticized, and so that alternatives can be considered. If the attempt is made to do science in accordance with falsificationism, bare or dressed, one substantial, influential and problematic assumption must remain implicit (as we have just seen), namely the metaphysical assumption that nature behaves as if simple or unified, no *ad hoc* theory being true. This is implicit in the adoption of the simplicity methodological principle of dressed falsificationism. AOE, by contrast, makes this implicit assumption explicit, and provides a framework within which rival versions can be proposed and critically assessed.

4. *Simplicity.* Falsificationism fails to say what the simplicity of a theory *is*. Bare falsificationism provides an account of simplicity in terms of falsifiability, but we have already seen that this account is untenable. Popper's (1963) "requirement of simplicity" appeals to a conception of simplicity or unity that is wholly in addition to falsifiability, but does not explain what the simplicity or unity of a theory is. It fails to explain how the simplicity of a theory can possibly be methodologically or epistemologically significant when a simple theory can always be made complex by a suitable

change of terminology, and *vice versa*. Popper himself recognized the inadequacy of his simplicity requirement when he called it "a bit vague", said that "it seems difficult to formulate it very clearly", and acknowledged that it threatened to involve one in infinite regress (Popper, 1963, p. 241). By contrast, AOE solves the problems of explaining what the simplicity or unity of a theory is without difficulty. The totality of fundamental physical theory, T , is unified to the extent that its *content* exemplifies physicalism. The more the content of T departs from exemplifying physicalism, the more disunified T is.¹⁰⁰ Because what matters is content, not form, the way T is *formulated* is irrelevant to this way of assessing simplicity or unity. Falsificationism cannot avail itself of this way of assessing unity because it involves acknowledging that physicalism is a basic tenet of scientific knowledge, something which falsificationism denies. Within AOE, there is a second way in which the unity of T may be assessed: in terms of the extent to which the content of T exemplifies the best available level 3 metaphysical blueprint. This second conception of simplicity or unity evolves with the evolution of level 3 ideas. As we improve our ideas about how the universe is unified, with the advance of knowledge in theoretical physics, so non-empirical methods for selecting theories on the basis of simplicity or unity improve as well.

Thus current symmetry principles of modern physics, such as Lorentz invariance and gauge invariance, which guide acceptance of theory, are an advance over simplicity criteria upheld by Newton. This account of simplicity can be extended to individual theories in two ways. First, we may treat an individual theory as a candidate theory of everything. Second, given two individual theories, T_1 and T_2 , and given the rest of fundamental theory, T , T_1 is simpler than T_2 iff $T + T_1$ is simpler than $T + T_2$, where the latter is assessed in one or other of the ways indicated above.¹⁰¹

¹⁰⁰. Dynamical theories are partially ordered with respect to the extent that they exemplify physicalism, with respect to their degree of unity, in other words. For further details see Maxwell (1998, chapter 4).

¹⁰¹. For a very much more detailed exposition of this solution of the problem of simplicity, together with an account of the way in which great unifying theories

It may be objected that this proposed solution to the problem of simplicity is circular: the unity of level 2 theory is explicated in terms of the unity of level 4 physicalism. But this objection is not valid. In order to solve the problem, it is not necessary to explicate what "simplicity" or "unity" mean; rather, what needs to be done is to show how theories can be partially ordered with respect to "simplicity" or "unity" in a way that does not depend on formulation. This is achieved by partially ordering theories in terms of how well their content exemplifies the content of physicalism, so that, roughly, the more the content of a theory violates the symmetries associated with the content of physicalism, the less unity it has. As long as physicalism is a meaningful thesis, and provides a formulation-independent way of partially ordering theories in the way indicated, this suffices to solve the problem. That physicalism embodies intuitive ideas of "unity" is a bonus. For a more detailed rebuttal of this objection, see Maxwell (1998, pp. 118-123; 2004, pp. 160-174).

5. *Scientific Method*. Dressed falsificationism acknowledge (correctly) that two considerations govern selection of theory in science, namely considerations that have to do with (a) evidence, and (b) simplicity. But because it cannot solve the problem of what simplicity *is*, dressed falsificationism cannot, with any precision, specify what methods are involved when theories are selected on the basis of simplicity. Nor can the view do justice to the way in which the methods of physics evolve with evolving knowledge, especially methods that assert that acceptable theories must satisfy this or that *symmetry*. In other words, falsificationism fails to solve what may be called the "methodological" problem of induction, the problem of specifying, merely, what the methods are that are employed by science in accepting and rejecting theories (leaving aside the further problem of justifying these methods given that the aim is to acquire knowledge). AOE, by contrast, solves the problem of simplicity, and thus can specify precisely what methods are involved when theories are selected on the basis of simplicity.

of physics illustrate the solution, see Maxwell (1998, chapters 3 and 4). See also Maxwell (2004, pp. 160-174).

Furthermore, AOE can do justice to *evolving* criteria of simplicity (as we have just seen), and hence evolving methods. According to AOE, the totality of fundamental physical theory, T, can be assessed with respect to how well its content exemplifies (i) the relatively fixed level 4 thesis of physicalism, or (ii) the evolving, best available level 3 thesis. Whereas (i) constitute fixed criteria of simplicity or unity (as long as physicalism is not modified), (ii) constitute evolving criteria, criteria of unity that improve with improving knowledge.

6. *Evolving Aims and Methods*. A point, briefly alluded to in 4 and 5 above, deserves further emphasis. As physics has evolved, from Newton's time to today, non-empirical methods, determining what theories will be accepted and rejected, have evolved as well. Newton, in his *Principia*, formulated four rules of reasoning, three of which are concerned with simplicity (Newton, 1962, vol 2, pp. 398-400). Principles that have been proposed since include: invariance with respect to position, orientation, time, uniform velocity, charge conjugation, parity, time-reversal; principles of conservation of mass, momentum, angular momentum, energy, charge; Lorentz invariance; Mach's principle, the principle of equivalence; principles of gauge invariance, global and local; supersymmetry; duality principles; the principle that different kinds of particle should be reduced to one kind, and different kinds of force should be reduced to one kind; the principle that space-time on the one hand, and particles-and-forces on the other, should be unified. All of these principles can be interpreted as methodological rules which specify requirements theories must meet in order to be accepted. They can also be interpreted as physical principles, making substantial assertions about such things as space, time, matter, force. Some, such as conservation of mass, parity, and charge conjugation, have been shown to be false; others, such as Mach's principle, have never been generally accepted; still others, such as supersymmetry, remain speculative.

Principles such as these, which can be interpreted either as physical assertions or as methodological principles, which are made explicit, developed, revised and, on occasions, rejected or refuted, are hard to account for within the framework of

falsificationism. It is especially difficult, within this framework, to account for principles which (a) have a quasi *a priori* role in specifying requirements theories must satisfy in order to be accepted, but which at the same time (b) make substantial *physical* assertions about the nature of the universe. AOE, on the other hand, predicts the existence of such principles, with just the features that have been indicated. Accepted principles are components of the currently accepted level 3 blueprint. As the accepted blueprint evolves, these principles, interpreted either as physical or methodological principles, evolve as well. Indeed, according to AOE, these principles, and associated blueprints, do not just evolve; they are *improved* with improving theoretical knowledge. AOE provides a more or less fixed framework of relatively unproblematic assumptions and associated methods (at level 4 or above) within which highly problematic level 3 assumptions and associated methods may be improved in the light of the empirical success and failure of rival research programmes (which adopt rival level 3 assumptions and associated methods).

This can be reformulated in terms of *aims* and methods of physics. A basic aim of theoretical physics is to discover the true theory of everything. This aim can be characterized in a range of ways, depending on how broadly or narrowly “theory of everything” is construed, what degree of unity such a theory must have in order to be a theory at all, and thus how much metaphysics is built into, or is presupposed by, the aim so characterized. The aim might be construed in such a way that no more than the truth of the thesis at level 7, or at level 6, is presupposed. Or, more specifically, the truth of the thesis at level 5 might be presupposed, or even more specifically, the truth of physicalism at level 4; or a range of increasingly specific blueprints at level 3 might be presupposed. Corresponding to these increasingly specific *aims* there are increasingly restrictive *methods*. As the aim becomes more specific, so it becomes more problematic, in that the presupposed metaphysics becomes increasingly likely to be false, which would make the corresponding aim unrealisable. AOE can thus be construed as providing a kind of nested framework of aims and methods, the aims becoming, as one goes down the hierarchy, increasingly problematic, and vulnerable to being unrealisable in

principle, because the presupposed metaphysics is false. Within the framework of relatively unspecific, unproblematic, permanent aims and methods (high up in the hierarchy) much more specific, problematic, fallible aims and methods (low down in the hierarchy) can be revised and improved in the light of improving knowledge. There is, as I have already in effect said, something like positive feedback between improving scientific knowledge and improving aims and methods. As knowledge improves, knowledge-about-how-to-improve-knowledge improves as well. This capacity of science to adapt itself - its aims and methods (its philosophy of science) - to what it finds out about the universe is, according to AOE, the methodological key to the astonishing progressive success of science. Falsificationism, with its fixed aim and fixed methods, is quite unable to do justice to this positive feedback, meta-methodological feature of science, this capacity of science to learn about learning as it proceeds.

7. *Verisimilitude*. The so-called problem of verisimilitude arises because physics usually proceeds from one false theory to another, thus rendering obscure what it can mean to say that science makes progress. Popper (1963, chapter 10 and Addenda) tried to solve this problem within the framework of falsificationism but, as Miller (1974) and Tichy (1974) have shown, this attempted solution does not work. Not only does falsificationism fail to specify properly the methods that make progress in theoretical physics possible; it fails even to say what progress in theoretical physics *means*.

AOE solves the problem without difficulty. First, the fact that physics does proceed from one false theory to another, far from undermining physicalism, and hence AOE as well, is just the way theoretical physics must proceed, granted physicalism (as I have already indicated). For, granted physicalism, any theory, T*, which captures precisely how phenomena evolve in some restricted domain, must be generalizable to cover *all* phenomena. If T* cannot be so generalized then, granted physicalism, it cannot be precisely true. In so far as physics proceeds by developing theories which apply to restricted, but successively increasing, domains of

phenomena, it is bound (granted physicalism) to proceed by proposing one false theory after another.

Second, AOE solves the problem of what it can *mean* to say that theories, T_0, \dots, T_N , get successively closer and closer to the true theory-of-everything, T , as follows. For this we require that T_N can be "approximately derived" from T (but not vice versa), T_{N-1} can be "approximately derived" from T_N (but not vice versa), and so on down to T_0 being "approximately derivable" from T_1 (but not vice versa).

The key notion of "approximate derivation" can be indicated by considering a particular example, the "approximate derivation" of Kepler's law that planets move in ellipses around the sun (K) from Newtonian theory (NT).

The "derivation" is done in three steps. *First*, NT is restricted to N body systems interacting by gravitation alone within some definite volume, no two bodies being closer than some given distance r . *Second*, keeping the mass of one object constant, we consider the paths followed by the other bodies as their masses tend to zero. According to NT , in the limit, these paths are precisely those specified by K for planets. In this way we recover the *form* of K from NT . *Third*, we reinterpret this "derived" version of K so that it is now taken to apply to systems like that of our solar system. (It is of course this *third* step of reinterpretation that introduces error: mutual gravitational attraction between planets, and between planets and the sun, ensure that the paths of planets, with masses greater than zero, must diverge, however slightly, from precise Keplerian orbits.)

Quite generally, we can say that T_{r-1} is "approximately derivable" from T_r if and only if a theory empirically equivalent to T_{r-1} can be extracted from T_r by taking finitely many steps of the above type, involving (a) restricting the range of application of a theory, (b) allowing some combination of variables of a theory to tend to zero, and (c) reinterpreting a theory so that it applies to a wider range of phenomena.

This solution to the problem of what progress in theoretical physics *means* requires AOE to be presupposed; it does not work if falsificationism is presupposed. This is because the solution requires one to assume (a) that the universe is such that a yet-to-be-

discovered, true theory of everything, T, exists, and (b) current theoretical knowledge can be approximately derived from T. Both assumptions, (a) and (b), are justified granted AOE; neither assumption is justifiable granted falsificationism.¹⁰²

8. *Discovery of New Fundamental Theories.* Given falsificationism, the discovery of new fundamental physical theories that turn out, subsequently, to meet with great empirical success, is inexplicable. (One thinks here of Newton's discovery of his mechanical theory and theory of gravitation, Maxwell's discovery of classical electromagnetism, Einstein's discovery of the special and general theories of relativity, Bohr's discovery of "old" quantum theory, Heisenberg's and Schrödinger's discovery of "new" quantum theory, Dirac's discovery of the relativistic quantum theory of the electron and, in more recent times, the discovery of quantum electrodynamics, the electroweak theory, quantum chromodynamics and the standard model.) Granted that a new theory is required to explain a range of phenomena, there are, on the face of it, infinitely many possibilities. In the absence of rational guidance towards good conjectures, it would seem to be infinitely improbable that anyone should, in a finite time, be able to come up with a theory that successfully predicts new phenomena. The only guidance that falsificationism can provide is to think up new theories that "proceed from some simple, new, and powerful, unifying idea", in accordance with Popper's (1963) requirement of simplicity, but this is so vague and ambiguous as to be almost useless. Famously, Popper explicitly denied that a rational method of discovery is possible at all: see Popper (1959, p. 31). But if discovery is not rational, it becomes miraculous that

¹⁰². It may be objected that if T is assumed to be the true *unified* theory of everything, no meaning can be given to the idea that theoretical physics is making progress, by means of a succession of false theories, to a more or less *disunified* theory of everything. But T does not need to be assumed to be unified; all that is required is that T is such that the notion of "partial derivation" from T makes sense. For further discussion of the inability of any standard empiricist view such as falsificationism to solve the problem of verisimilitude, and the ability to AOE to solve the problem, see Maxwell (1998, pp. 70-72, 211-217 and 226-227).

good new theories are ever discovered. Scientific progress becomes all but inexplicable.

AOE, by contrast, provides physics with a rational, if fallible and non-mechanical, method for the discovery of new fundamental physical theories. This method involves modifying the current best level 3 blueprint so that:

(a) the new blueprint exemplifies physicalism better than its predecessor;

(b) the new blueprint promises, when made sufficiently precise to become a testable theory, to unify clashes between predecessor theories;

(c) the new theory promises to exemplify the new blueprint better than the predecessor theories exemplify the predecessor blueprint.

(a), (b) and (c) provide means for assessing how good an idea for a new theory is which do not involve empirical testing (which is brought in once the new theory has been formulated). The level 4 thesis of physicalism provides continuity between the state of knowledge before the discovery of the new theory, and the state of knowledge after this discovery. Modifying the current level 3 blueprint ensures that the new theory will be incompatible with its predecessors; it will postulate new kinds of entities, forces, space-time structure, and will exhibit new symmetries. In other words, because of the hierarchical structure of AOE, there is (across revolutions) both continuity (at level 4) and discontinuity (at levels 2 and 3), something that is not possible given falsificationism. AOE provides physics with specific non-empirical tasks to perform, specific non-empirical problems to be solved, and non-empirical methods for the assessment of ideas for new theories, all of which adds up to a rational, if fallible, method of discovery. It all stems from recognizing that physicalism is a part of current scientific knowledge. The discovery of new fundamental physical theories ceases to be inexplicable. None of this is possible granted falsificationism.¹⁰³

¹⁰³. For further discussion of the method of discovery provided by AOE see Maxwell (1974, Part II; (1993, Part III); and (1998, pp. 159-163 and 219-223).

The fact that AOE is able to provide a rational method of discovery, while falsificationism is not, is due to the greater rigour of AOE (a point mentioned in 3 above). AOE has greater rigour because AOE acknowledges, while falsificationism denies, metaphysical assumptions implicit in persistent scientific preference for simple, explanatory theories. It is precisely the explicit acknowledgement of these metaphysical assumptions which makes the rational method of discovery of AOE possible.

9. *Diversity of Scientific Method.* One striking feature of natural science, often commented on, is that different branches of the natural sciences have somewhat different methods. Experimental and observational methods, and methods or principles employed in constructing and assessing theories, vary as one moves from theoretical to phenomenological physics, from physics to chemistry, from astronomy to biology, from geology to ethology. Falsificationism can hardly do justice to this striking diversity of method within the natural sciences. Popper, indeed, tends to argue that there is unity of method, not only in natural science, but across the whole of science, including social science as well: see Popper (1961). AOE, by contrast, predicts diversity of method throughout natural science, overlaid by unity of method at a meta-methodological level. AOE can do justice to the diversity of methods to be found in diverse sciences, without underlying unity and rationality being sacrificed.

It is important to appreciate, first, that different branches of the natural sciences are not isolated from one another: they form an interconnected whole, from theoretical physics to molecular biology, neurology and the study of animal behaviour. Different branches of natural science, even different branches of a single science such as physics, chemistry or biology, have, at some level of specificity, different aims, and hence different methods. But at some level of generality all these branches of natural science have a common aim, and therefore common methods: to improve knowledge and understanding of the natural world. All (more or less explicitly) put AOE into practice, but because different scientific specialities have different specific aims, at the lower end of the hierarchy of methods different specialities have somewhat

different methods, even though some more general methods are common to all the sciences. Furthermore, all natural sciences apart from theoretical physics presuppose and use results from other scientific specialities, as when chemistry presupposes atomic theory and quantum theory, and biology presupposes chemistry. The results of one science become a part of the presuppositions of another, implicit in the aims of the other science (equivalent to the level 3 blueprint of physics, or the level 4 thesis of physicalism). This further enhances unity throughout diversity, and helps explain the need for diversity of method.

But in order to exhibit the rationality of the diversity of method in natural science, apparent in the evolution of methods of a single science, and apparent as one moves, at a given time, from one scientific speciality to another, it is essential to adopt the meta-methodological, hierarchical standpoint of AOE, which alone enables one to depict methodological unity (high up in the hierarchy) throughout methodological diversity (low down in the hierarchy). Falsificationism, lacking this hierarchical structure, cannot begin to do justice to this key feature of scientific method, diversity at one level, unity at another; nor can it begin to do justice to the rational *need* for this feature of scientific method.

There is a further, important point. Any new conception of science which improves our understanding of science ought to enable us to improve scientific practice. It would be very odd if our ability to do science well were wholly divorced from our understanding of what we are doing. A test for a new theory of scientific method ought to be, then, that it improves scientific practice, and does not merely accurately depict current practice. AOE passes this test. In providing a framework for the articulation and scrutiny of level 3 metaphysical blueprints, as an integral part of science itself, thus providing a rational means for the development of new non-empirical methods, new symmetry principles, and new theories, AOE advocates, in effect, that current practice in theoretical physics be modified. AOE makes explicit what is at present only implicit. And more generally, in depicting scientific method in a hierarchical, meta-methodological fashion, AOE has implications for method throughout the natural sciences, and not just for theoretical physics.

In case it should seem miraculous that science has made progress without AOE being generally understood and accepted, I should add that good science has always put something close to AOE into practice in an implicit, somewhat covert way, and it is this which has made progress possible.

7 Thomas Kuhn

As I remarked in section 1 above, the main difference between Kuhn's (1962, 1970a) picture of science and Popper's is that, whereas Kuhn stresses that, within normal science, paradigms are dogmatically protected from refutation, from criticism, Popper holds that theories must always be subjected to severe attempted refutation. AOE is even more Popperian than Popper's falsificationism, in that AOE exposes to criticism assumptions that falsificationism denies, and thus shields from criticism. One might think, therefore, that AOE would differ even more from Kuhn's picture of science than falsificationism does.

It is therefore rather surprising that exactly the opposite is the case. In some important respects, AOE is closer to Kuhn than to Popper.

The picture of science that emerges from Kuhn (1970a) may be summarized like this. There are three stages to consider. First, there is a pre-scientific stage: the discipline is split into a number of competing schools of thought which give different answers to fundamental questions. There is debate about fundamental questions between the schools, but no overall progress, and no science.

Second, the ideas of one such school begin to meet with empirical success; these ideas become a "paradigm", and the pre-scientific school becomes normal science (competing schools withering away). Within normal science, no attempt is made to refute the paradigm (roughly, the basic theory of the science); indeed, the paradigm may be accepted even though there are well known apparent refutations. When the paradigm fails to predict some phenomenon, it is not the paradigm, but the skill of the scientist, that is put to the test. The task of the normal scientist is to solve puzzles, rather than problems. The paradigm specifies what is to count as a solution, specifies what methods are to be

employed in order to obtain the solution, and guarantees that the solution exists: these are all characteristics of puzzles rather than open-ended problems. The task is gradually to extend the range of application of the paradigm to new phenomena, textbook successes being taken as models of how to proceed. Methods devolve from paradigms.

Third, the paradigm begins to accumulate serious failures of prediction; these resist all attempts at resolution, and some scientists lose faith in the capacity of the paradigm to overcome these "anomalies". A new paradigm is proposed, which does resolve these recalcitrant anomalies, but which may not, initially, successfully predict all that the old paradigm predicted. Empirical considerations do not declare that the new paradigm is, unequivocally, better than the old. Normal science gives way to a period of revolutionary science. Scientists again debate fundamentals, arguments for and against the rival paradigms often presupposing what they seek to establish. Rationality breaks down. If the revolution is successful, the new paradigm wins out, and becomes the basis for a new phase of normal science. Many old scientists do not accept the new paradigm; they die holding onto their convictions.

Kuhn argues that the dogmatic attitude inherent in normal science is necessary if science is to make progress. Applying a paradigm to new phenomena, or to old phenomena with increasing accuracy, is often extremely difficult. If every failure was interpreted as a failure of the paradigm, rather than of the scientist, paradigms would be rejected before their full range of successful application had been discovered. By refusing to reject a paradigm until the limits of its successes have been reached, scientists put themselves into a much better position to develop and apply a new paradigm. For reasons such as these, normal science, despite being ostensibly designed to discover only the expected, is actually uniquely effective in disclosing novelty. Popper (1970), in criticizing Kuhn on normal science, ignored these arguments in support of the necessity of normal science for scientific progress.

AOE holds that much scientific work ought indeed to resemble Kuhn's normal science, in part for reasons just indicated. But there are even more important considerations. According to AOE, and in

sharp contrast with falsificationism, theoretical physics accepts a level 3 metaphysical blueprint, which exercises a powerful constraint on what kind of new theories physicists can try to develop, consider or accept. The blueprint has a role reminiscent, in some respects, of Kuhn's paradigm, and theoretical physics, working within the constraints of the blueprint, its non-empirical methods set by the blueprint, has some features of Kuhn's normal science.

Furthermore, according to AOE, other branches of natural science less fundamental than theoretical physics invariably presuppose relevant parts of more fundamental branches. Thus chemistry presupposes relevant parts of atomic theory and quantum theory; biology relevant parts of chemistry; astronomy relevant parts of physics. Such presuppositions of a science have a role, for that science, that is analogous to the role that the current level 3 blueprint, or the level 4 thesis of physicalism, has for theoretical physics. The presuppositions act as a powerful constraint on theorizing within the science. They set non-empirical methods for that science. Such presuppositions have a role, in other words, which is similar, in important respects, to Kuhn's paradigms. Viewed from an AOE perspective, one can readily see how and why much of science is Kuhnian puzzle-solving rather than Popperian problem-solving.

There are also, it must be emphasized, major differences between Kuhn and AOE. The chief difference is that, according to AOE, science has a paradigm for paradigms - to put it in Kuhnian terms. In order to be acceptable, level 3 blueprints must exemplify the level 4 thesis of physicalism (which in turn must exemplify the level 5 thesis of comprehensibility and so on, up to level 7). This means that, as long as physicalism continues to be accepted as the best available level 4 thesis for science, metaphysical blueprints can be assessed in a quasi non-empirical way, in terms of how well they accord with physicalism. Natural science is, according to AOE, one sustained, gigantic chunk of normal science, with physicalism as its paradigm. In this respect, AOE is more Kuhnian than Kuhn (in addition to being more Popperian than Popper!).

Like falsificationism, Kuhn's picture of science is hardly tenable. In the first place, it does not fit scientific practice very well.

Normal science undoubtedly exists, as even Popper recognized; it may well be that most scientific activity has the character of Kuhn's normal science. But even when a discipline seems most like normal science, almost always there are a few scientists actively engaged in developing alternatives to the reigning paradigm. And on occasions, it is from the work of these few that a new paradigm, and a new phase of normal science springs, often in a way that is quite different from Kuhn's account. It is not obvious that accumulation of anomalies, resulting in a crisis in biology, led to Darwin's theory of evolution. Quantum theory did not emerge, initially, from a crisis in classical physics. Planck's work around 1900 on black body radiation engendered the quantum revolution. It is true that classical physics, applied to a so-called black body emitting electromagnetic radiation, made a drastically incorrect prediction, but no one, not even Planck, thought that this posed a serious problem for classical physics. The fallacious prediction of classical physics was dubbed "the ultra-violet catastrophe"; but this phrase was coined by Ehrenfest, after the quantum revolution was under way, around 1912, as propaganda for the new theory. It was Einstein who first recognized that Planck's work spelled the downfall of classical physics; but general recognition of this only came later, probably with Bohr's quantum theory of the atom, around 1913. Again, Einstein's general theory of relativity emerged, not because Newton's theory had accumulated anomalies and was in a state of crisis, but because it contradicted special relativity. Einstein sought a theory of gravitation compatible with special relativity, and it was this that led him to general relativity. These three revolutions, resulting in Darwinian theory, quantum theory and general relativity, are among the biggest and most important in the history of science; and yet they do not fit Kuhn's pattern.

Failure to fit scientific practice in detail does not, however, provide decisive grounds for rejecting a normative account of scientific method. One can always reply that the account specifies how science *ought* to proceed, not how it has in fact proceeded. Much more serious are the objections of principle to Kuhn's account. Kuhn, like Popper, provides no account of the creation of new paradigms. And given Kuhn's insistence that a new paradigm,

after a successful revolution, is incommensurable with its pre-revolutionary predecessor, it would seem impossible to provide rational (if fallible) procedures for the creation of good new paradigms while maintaining consistency with the rest of Kuhn's views. Kuhn does allow that non-empirical criteria, or values, such as consistency and simplicity, are employed by science permanently (and therefore, presumably, across revolutions) to assess theories or paradigms; but Kuhn also emphasizes that these criteria are flexible, and open to different interpretations (Kuhn, 1970a, p. 155; 1977, ch. 13). There is no account of what simplicity is, and no advance over Popper's "requirement of simplicity". Furthermore, Kuhn's appeal to simplicity faces the same difficulty we have seen arising in connection with Popper's appeal to simplicity. If "simplicity" is interpreted in such a way that it has real content, and is capable of excluding "complex" or disunified and aberrant theories or paradigms from science, then its permanent employment by science commits science to a permanent metaphysical assumption that persists through revolutions, something Kuhn explicitly rejects (and could not, in any case, provide a rationale for). If "simplicity" is interpreted sufficiently loosely and flexibly to ensure that no such metaphysical thesis is involved, invoking simplicity must fail to exclude complex, disunified, aberrant paradigms from science. Any Kuhnian requirement of simplicity, in short, must either be incompatible with the rest of Kuhn's views, or toothless and without content. Either way, Kuhn has no consistent method for excluding complex, aberrant paradigms from consideration. It should be noted that Kuhn is emphatic that no sense can be made of the idea that there is progress in knowledge across revolutions, the new paradigm being better, closer to the truth, than the old one: see Kuhn (1970a, chapter XIII). But this is a disaster for Kuhn's whole view. Why engage in normal science if the end result is the rejection of all that has been achieved, all the progress in knowledge of that period of normal science being sacrificed when the science adopts a new paradigm? Kuhn's arguments for the progressive character of normal science, indicated above, are all defeated.

Perhaps the most serious objection to Kuhn's picture of science is the obvious basic unintelligence of its prescriptions for scientific

research. Suppose we have the task of crossing on foot difficult terrain, containing ravines, cliffs, rivers, swamps, thickets. Kuhn's view, applied to this task, would be as follows. After debate about which route to follow (pre-science), one particular route is chosen and then followed with head down, no further consideration being given to changing the route (normal science). Eventually, this leads to an impasse: one comes face to face with an unclimbable cliff, or finds oneself waist deep in a swamp, and in danger of drowning (crisis). Finding oneself in these dire circumstances, a new route is taken (new paradigm), and again, with head down, this new route is blindly followed (normal science) until, again, one finds oneself unable to proceed, about to drown in a river, or tumble into a ravine.

This is clearly a stupid way to proceed. It would be rather more intelligent if, as one tackles immediate problems of wading through this stream, climbing down this scree (puzzle-solving of normal science), one looks ahead, whenever possible, and reconsiders, in the light of the terrain that has been crossed, what adjustments one needs to make to the route one has opted to follow. Exactly the same point holds for science. There can be division of labour. Even if a majority of scientists tackle the multitude of puzzles that go to make up normal scientific research, taking the current theory, or paradigm, for granted, there ought also to be some scientists who are concerned to look ahead, consider more fundamental problems, explore alternatives to the current paradigm. In this way new paradigms may be developed before science plunges deep into crisis. And just this does go on in scientific practice, as I have already indicated in the brief discussion of the work of Darwin and Einstein (and somewhat less convincingly, Planck). Another example of a new, revolutionary theory or paradigm being proposed in the absence of crisis is Wegener's advocacy of the movement of continents, anticipating the plate tectonic revolution by decades. Science is, in practice, more intelligent than Kuhn allows.

In sharp contrast to Kuhn, AOE does not merely stress the importance of "looking ahead", of trying to develop new theories, new paradigms before science has plunged into crisis; even more important, AOE provides a framework for theoretical physics (and

therefore, in a sense, for the whole of natural science) within which ideas for fundamental new theories may be developed and assessed.

According to Kuhn, successful revolutions mark radical discontinuities in the advancement of science, to the extent, indeed, that old and new paradigms are "incommensurable" (i.e. so different that they cannot be compared). This Kuhnian view is most likely to be correct when applied to revolutions in fundamental theoretical physics, where radical discontinuity seems most marked. But it is precisely here that Kuhn's claim turns out to be seriously inadequate. As I have already emphasized, all revolutions in theoretical physics, despite their diversity in other respects, reveal one common theme: they are all gigantic steps in unification. From Newton, via Maxwell, Einstein, Bohr, Schrödinger and Dirac, to Salam, Weinberg and Gell-Mann, all new revolutionary theories in physics bring greater unity to physics. (And Darwinian theory, one might add, brings a kind of unity to the whole of biology.) The very phenomenon that Kuhn holds to mark discontinuity, namely revolution, actually also reveals continuity - continuity of the search for, and the successful discovery of, underlying theoretical unity.

This aspect of natural science, to which Kuhn fails entirely to do justice, is especially emphasized by AOE. According to AOE, revolutions in theoretical physics mark discontinuity at the level of theory, at level 2, and even discontinuity at level 3, but continuity at level 4. Physicalism, which asserts that underlying dynamic unity exists in nature, persists through revolutions - or, at least, has persisted through all revolutions in physics since Galileo. In order to make rational sense of natural science, we need to interpret the whole enterprise as seeking to turn physicalism, the assertion of underlying dynamic unity in nature, into a precise, unified, testable, physical "theory of everything". That, in a sentence, is what AOE asserts. Physicalism, according to AOE, despite its metaphysical (untestable) character, is the most secure item of theoretical knowledge in science; it is *the* most fruitful idea that science has come up with, at that level in the hierarchy of assumptions.

Because of its recognition that, despite the discontinuity of revolutions at levels 2 and 3, there is the continuity of the persistence of physicalism at level 4 (and of other theses at levels higher up in the hierarchy), AOE is able to resolve problems concerning the discovery and assessment of paradigms which Kuhn's view is quite unable to solve. Both fundamental physical theories, and level 3 blueprints, can be partially ordered with respect to how well they exemplify physicalism, entirely independent of ordinary empirical assessment. Assessing progress through revolution poses no problem for AOE. As we have seen, AOE solves the problem of verisimilitude.

I have already mentioned that AOE does not merely describe scientific practice; it carries implications as to how scientific practice can be improved. One such implication concerns scientific revolutions. Kuhn (1970a) gives a brilliant description of the way, during a scientific revolution, there is a breakdown of rationality, competing arguments for the rival paradigms being circular, each presupposing what is being argued for. This is a feature of actual science. Scientists do not know how to assess competing theories objectively, when empirical considerations are inconclusive. But all this can be seen to be a direct consequence of trying to do science without persisting metaphysical assumptions concerning the comprehensibility of the universe, there thus being nothing available to constrain acceptance of theories when empirical considerations are inconclusive. Consider Kuhn's breakdown of rationality. A substantial revolution will involve, not just two rival paradigms or theories, T_1 and T_2 , but two rival blueprints, B_1 lurking behind T_1 , and B_2 lurking behind T_2 . Granted B_1 , T_1 is far more acceptable than T_2 , but the reverse granted B_2 . But B_1 and B_2 , being untestable, metaphysical theses, are not explicitly discussable, and objectively assessable, within science: so they are more or less repressed, excluded from discussion. Nevertheless, scientists do think in terms of B_1 and B_2 . Kuhn's Gestalt switch, involved in switching allegiance from T_1 to T_2 , can be pin-pointed as the act of abandoning the old blueprint and adopting the new one. Non-empirical arguments in favour of T_1 or T_2 can only take the form of an appeal to B_1 or B_2 , in however a muffled a way (due to the point that blueprints are not open to explicit discussion).

Such arguments will be circular, and entirely unconvincing to the opposition, in just the way described by Kuhn. Accept B_1 , and T_1 becomes the only possible choice; accept B_2 and T_2 is the only choice. Each side in the dispute is convinced that the other side is wrong, even incoherent. What needs to be done, and cannot be done, of course, is to discuss the relative merits of B_1 and B_2 . Just this can be done, granted AOE. T_1 , B_1 , T_2 and B_2 can all be assessed from the standpoint of adequacy in exemplifying physicalism. When the scientific community adopts AOE, the Kuhnian irrationality of revolutions will disappear from science.

It may be asked: How is it possible for AOE to be *both* more Popperian than Popper, and more Kuhnian than Kuhn? The answer is that AOE is more Popperian than Popper in making explicit, and so criticizable, metaphysical theses which falsificationism denies, and thus leaves implicit and uncriticizable within science. But AOE is also more Popperian than Popper in insisting we need to exploit criticism *critically*, so that it furthers, and does not sabotage, the growth of knowledge. Criticism needs to be marshalled and directed at that part of our conjectural knowledge which it is, we conjecture, the most fruitful to criticize. This means directing critical fire at level 2 theories and level 3 blueprints, it being less likely, though still possible, that criticism of the level 4 thesis of physicalism will aid the growth of empirical knowledge. Physicalism has played an extraordinarily fruitful role in the advancement of scientific knowledge; it should not be abandoned unless an even more apparently fruitful idea is forthcoming, or unless the empirical and explanatory success that physicalism appears to have engendered turns out to be illusory. It is this persistence of physicalism, for good Popperian reasons, which gives to theoretical physics, and indeed to the whole of natural science, something of the character of Kuhn's normal science, with physicalism as its quasi-permanent "paradigm".

8 Imre Lakatos

Lakatos sought to reconcile the very different views of science held by Popper and Kuhn. According to Kuhn, far from seeking falsifications of the best available theory, as Popper held, scientists protect the accepted theory, or "paradigm", from refutation for

most of the time, the task being to fit recalcitrant phenomena into the framework of the paradigm. Only when refutations become overwhelming, does crisis set in; a new paradigm is sought for and found, a revolution occurs, and scientists return to doing "normal science", to the task of reconciling recalcitrant phenomena with the new paradigm. Lakatos sought to reconcile Popper and Kuhn by arguing that science consists of competing fragments of Kuhnian normal science, or "research programmes", to be assessed, eventually, in terms of their relative empirical success and failure. Instead of research programmes running in series, one after the other, as Kuhn thought, research programmes run in parallel, in competition, this doing justice to Popper's demand that there should be competition between theories (a point emphasized especially by Feyerabend).¹⁰⁴ Lakatos became so impressed with the Kuhnian point that theories always face refutations, the empirical successes of a theory being a far more important guide to scientific progress than refutation, that he finally came to the conclusion that Popper's philosophy of science was untenable.

AOE has a number of features in common with Lakatos's methodology of scientific research programmes. AOE makes extensive use of the notion of scientific research programme. Like Lakatos's view, AOE exploits the idea that such research programmes can, sometimes, be compared with respect to how empirically progressive they are. AOE, again like Lakatos's view, sees the whole of science as a gigantic scientific research programme. And like Lakatos's view, AOE can be construed as synthesizing Popper's and Kuhn's views.

But there are also striking differences. There are differences in the way scientific research programmes are conceived, especially research programmes in fundamental physics. For Lakatos, main components of a research programme are the "hard core" (corresponding to Kuhn's "paradigm"), and the "protective belt" of "auxiliary hypotheses", which facilitate the application of the hard core to empirical phenomena. The main business of a research programme is to develop the protective belt, thus extending, and

¹⁰⁴. See Lakatos (1970, 1978). For Feyerabend's argument that severe testing requires the development of rival theories see Feyerabend (1965).

making more accurate, the empirical predictions of the hard core. The hard core is a testable theory rendered metaphysical by the methodological decision not to allow it to be refuted, refutations being directed at the protective belt rather than the hard core.

According to AOE, by contrast, the metaphysical kernel of a research programme is not a testable theory but rather a thesis that is genuinely metaphysical (i.e. more or less unspecific, and usually untestable) - a thesis such as the corpuscular hypothesis, Boscovich's point-atom blueprint, Einstein's unified field blueprint, and so on. The basic aim of the programme is to turn the relatively unspecific blueprint into a precise, testable (and true) physical theory. The research programme thus consists of a succession of theories, T_1, T_2, \dots, T_n , which can be compared, not only with respect to empirical success, but also with respect to how adequately each theory encapsulates, or exemplifies, the blueprint of the programme. (The latter is not possible within a Lakatosian programme.) Whereas a Lakatosian programme has a fixed basic theory (or hard core), and seeks to improve auxiliary hypotheses (the protective belt), an AOE programme strives to capture the blueprint more and more adequately by means of testable physical theories.

Both Lakatos's view and AOE permit one to see natural science as one gigantic research programme, but how this programme is construed is very different. For Lakatos "science as a whole can be regarded as a huge research programme with Popper's supreme heuristic rule: 'devise conjectures with more empirical content than their predecessors'" (1970, p. 132). The huge research programme of natural science has, for Lakatos, no hard core; to this extent, Lakatos's view is a variant of Popper's.¹⁰⁵ According to AOE,

¹⁰⁵. Granted Lakatos's overall view, the research programme of science cannot have a hard core, for then, in order to ensure Popperian severe testing, there would need to be a rival research programme with a rival hard core – and that would mean the original research programme was not the whole of science. Actually, Lakatos is not quite consistent here; after the sentence quoted in the text, Lakatos goes on "Such methodological rules may be formulated, as Popper has pointed out, as metaphysical principles. For instance, the *universal* anti-conventionalist rule against exception-barring may be stated as the metaphysical principle: 'Nature does not allow exceptions'" (1970, p. 132). That this admission

however, if natural science is viewed as one gigantic research programme, then it does have something like a hard core. First, there is physicalism at level 4, a metaphysical but nevertheless substantial thesis about the nature of the universe. And then there is the current blueprint at level 3, an even more substantial metaphysical thesis about the nature of the universe. These provide severe constraints on what theories are acceptable that are not straightforwardly empirical,¹⁰⁶ something that is not possible given the views of Popper or Lakatos¹⁰⁷ (or even Kuhn).

Lakatos and AOE have very different motivations for taking scientific research programmes so seriously. For Lakatos, the motivation comes from appreciating that a scientific theory, T, cannot be decisively refuted at an instant, as it were, partly because auxiliary hypotheses can always be invented to salvage T from a refutation, partly because early applications of a new theory, such as Newton's, may make simplifying assumptions which may well lead to false predictions (the fault lying with the simplifying, auxiliary hypotheses rather than the basic theory). Only by looking at a series of theories, a given T₁ (the hard core) plus changing auxiliary hypotheses (the protective belt), and comparing this with a rival series based on a different hard core, T₂, and comparing the extent to which the two series are empirically progressive or degenerating, can one assess the relative empirical merits of T₁ and T₂. For AOE, the situation is very different. A research programme

is damaging for Popper's bare falsificationism was pointed out in footnote 6; it is equally damaging for Lakatos's version of Popperianism.

¹⁰⁶. I say "not straightforwardly empirical" because both physicalism and the best available blueprint are themselves accepted on the grounds that they support a more empirically progressive research programme than any rival theses. Long-term empirical considerations influence choice of theses at levels 3 and 4, while at the same time these theses can lead to the rejection of potentially empirically successful theories that clash too severely with them (i.e. are too severely *ad hoc*).

¹⁰⁷. The Popperian and Lakatosian demand that theories be strictly universal places weak but rigid constraints on what theories are acceptable; the demand of AOE that theories accord, as far as possible, with physicalism and the best available blueprint, places strong, but flexible and revisable constraints on what theories are acceptable. For further discussion see Maxwell (1998, pp. 89-102, chapter 4, and 223-227.)

in theoretical physics consists of a blueprint, B, and a succession of theories, $T_1, T_2 \dots T_n$ (each equivalent to a Lakatosian hard core), successive attempts to capture B as a testable theory. If $T_1, T_2 \dots T_n$ are increasingly empirically successful (in a roughly Popperian sense) and also increasingly successful at capturing B, then this means that B is empirically fruitful. A rival blueprint, B^* , might be such that the series $T_1, T_2 \dots T_n$ moves further and further away from B^* : this would mean that B^* is empirically sterile. A major part of the point of research programmes, for AOE, is to assess the relative empirical fruitfulness of rival metaphysical theses, at levels 3 and 4 (and above, if necessary). Though mostly untestable, nevertheless metaphysical theses can be assessed in a quasi-empirical way, in terms of the empirical progressiveness or degeneracy of the research programmes with which they are associated (or can be regarded as being associated).¹⁰⁸ This is, according to AOE, a key feature of scientific method, one which makes scientific progress possible. It makes it possible for improving theoretical knowledge to lead to a reassessment of what is the best available blueprint, which in turn leads to a reassessment of the best available non-empirical methodological rules, such as symmetry principles. In other words, it makes it possible for there to be positive feedback between improving knowledge and improving aims-and-methods (improving knowledge-about-how-to-improve-knowledge), a vital feature of scientific rationality according to AOE.

The differences indicated enable AOE to overcome problems which Lakatos's view cannot solve. Lakatos insists that there is no such thing as instant rationality: however apparently decisive the refutation of a theory may be, it is always possible to salvage it from refutation in a content increasing way by the invention of an appropriate auxiliary hypothesis. It is this consideration which leads Lakatos to argue that only series of theories, competing research programmes, can be assessed rationally, in terms of relative empirical progressiveness. But in practice in science there do seem to be instant refutations. A famous example is the refutation of parity. This is a symmetry which declares, roughly,

¹⁰⁸. For further details and discussion, see Maxwell (1998, pp. 172-180).

that if a process can occur, then so can its mirror image. This was decisively refuted by Wu et al. (1957), by means of an experiment which showed that electrons were emitted in a preferential direction from cobalt nuclei undergoing radioactive decay in a magnetic field. Parity conservation implied that this would not occur. Strictly speaking, it was not parity conservation on its own that was refuted, but parity plus quantum theory plus the theory of weak interactions plus the theory of nuclear structure plus a highly theoretical description of the experiment. One would think there was plenty of scope, here, for auxiliary hypotheses to be invented to salvage parity from refutation. No such hypothesis was forthcoming; the refutation of parity conservation was accepted immediately by the physics community, despite strong resistance to accepting such a conclusion (because of the implausibility of supposing that nature distinguishes between left-handedness and right-handedness at the level of fundamental physical theory). Allan Franklin, who has produced what is probably the best account of the downfall of parity conservation, has put the matter like this: "It is fair to say that as soon as any physicist saw the experimental result they were convinced that parity was not conserved in the weak interactions" (Franklin, 1990, p. 66).¹⁰⁹ Scientific practice seems almost to refute Lakatos's view.

But it does not refute AOE. According to Lakatos, in the end only empirical considerations, plus considerations of empirical content, restrict choice of theory; few restrictions are placed on how a body of theory may be modified to salvage it from refutation. AOE places much more severe restrictions on choice of theory. In addition to those that it has in common with Lakatos's view, AOE demands of a fundamental physical theory that it, together with other such theories, exemplifies physicalism, to a sufficient degree. This makes it very much more difficult to modify a body of theory so as to salvage it from refutation. Instant refutation is not surprising, granted AOE.

¹⁰⁹. For an account of the discovery of parity non-conservation, and of the decisive character of the experiments refuting parity conservation, see Franklin (1990, pp. 63-6 and 151-2). See also Franklin (1986).

Lakatos's view requires that science consists of competing research programmes. Unquestionably, the history of science reveals that competing research programmes have, on occasions existed. But it is not clear that all science has this character, as Lakatos's view would seem to require. After Heisenberg and Schrödinger had developed quantum theory in the mid 1920's, there continued to be debate about how the new theory should be interpreted, and whether the new theory, interpreted along the orthodox lines advocated by Bohr, Heisenberg and others, was ultimately acceptable. But there was nothing like a competing research programme. Viewed from the perspective of AOE, all this makes perfect sense. There were indeed serious grounds for regarding the new theory as unsatisfactory (see Maxwell, 1998, chapter 7). But the new theory had achieved such striking successes, it was rational to conjecture that progress lay in developing the new theory, applying it to new phenomena, reconciling it with special relativity - in doing something like Kuhnian normal science, in other words - rather than in trying to develop a rival theory, a rival research programme. (To say this is not to say that serious attention should not have been given to the theoretical defects of orthodox quantum theory.) Not only does the history of science fail to reveal that there are always competing research programmes; whenever a new theory arrives on the scene that meets with extraordinary empirical success and no refutation, no good rationale may exist for inventing a rival research programme. (As we have seen, unlike Popper's falsificationism and Lakatos's research programme view, AOE holds that something like Kuhn's normal science may well be rational, as long as it is accompanied by some sustained tackling of problems associated with the currently accepted blueprint. This may, eventually, but not immediately, lead to the development of a new fundamental theory, a new research programme.)

There are other, much more decisive ways in which AOE is an improvement over Lakatos's view. Lakatos's methodology of research programmes inherits a number of unsolved problems from its two sources, Popper and Kuhn. Like Popper and Kuhn, Lakatos has no solution to the problem of what the simplicity, unity or explanatory character of a theory, or hard core, is; AOE, as I have

indicated briefly above, solves the problem without difficulty. In failing to say what simplicity *is*, Lakatos also fails to articulate with any precision that part of scientific method concerned with simplicity; AOE faces no difficulty here either. Like Popper and Kuhn, Lakatos can say nothing useful about how new theories, new hard cores, are created or discovered; AOE, as a result of including levels 3 and 4 within the domain of scientific knowledge, is able to specify a rational, if fallible and non-mechanical, method for the creation of new theories, even new fundamental theories of physics. Finally, Lakatos's view fails to solve the problem of verisimilitude, a problem which can be readily solved granted AOE.

Popper, Kuhn and Lakatos, despite their differences, have one big failure in common (the source of almost all the others). All three take for granted that:

(A) In science no untestable but nevertheless substantial thesis about the world can be accepted as a part of scientific knowledge in such a firm way that theories which clash with it, even if highly successful empirically, are nevertheless rejected.

Popper accepts (A) in that, for him, untestable theses are metaphysical, and therefore not a part of scientific knowledge. Kuhn holds it, because, for Kuhn, nothing theoretical survives a revolution. Kuhn's acceptance of (A) is also apparent in his whole treatment of revolutions: precisely because Kuhn accepts (A), Kuhn cannot invoke anything like the level 4 thesis of physicalism to assess rival paradigms during a revolution, when empirical considerations are inconclusive. The Kuhnian irrationality of revolutions is a consequence of scientists accepting (A); and in so far as Kuhn thinks this irrationality is inevitable, Kuhn accepts (A) as well.

A case could be made out for saying that Lakatos came near to rejecting (A) in arguing for the need for science to adopt a conjectural metaphysical inductive principle which, if true, would more or less guarantee that Popperian, or rather Lakatosian, methods deliver authentic theoretical knowledge.

But Lakatos here missed the fundamental point, central to AOE, and highly Popperian in spirit, that our current methods are all too likely to be more or less the *wrong* methods to adopt, the

metaphysics implicit in these methods being *false*, there thus being a vital need, for scientific progress, to make the metaphysics explicit so that it can be criticized, so that alternatives can be developed and considered, leading to *improved* metaphysics and methods, this in turn requiring the development of a hierarchy of metaphysical theses to form a framework of relatively unproblematic theses within which more specific problematic theses may be developed and assessed.

Interestingly enough, Lakatos himself was aware of this deficiency in his "plea to Popper for a whiff of 'inductivism'" (1978, p. 159). Discussing his proposal that one should appeal to a metaphysical inductive principle as a conjecture as a part of the solution to the problem of induction, Lakatos says:

"Alas, a solution is interesting only if it is embedded in, or leads to, a major research programme; if it creates new problems - and solutions - in turn. *But this would be the case only if such an inductive principle could be sufficiently richly formulated so that one may, say criticize our scientific game from its point of view.* My inductive principle tries to explain why we 'play' the game of science. But it does so in an *ad hoc*, not in a 'fact-correcting (or, if you wish, 'basic value judgment correcting') way" (Lakatos, 1978, p. 164).

Lakatos highlights, here, the difference between his own position and that of AOE. The (revisable) AOE thesis of physicalism is indeed "sufficiently richly formulated so that one may...criticize our scientific game from its point of view". AOE not only offers a new research programme for the philosophy of science; it modifies the research programme of science, one modification being that the philosophy of science becomes an integral part of science itself. The passage above makes me wonder whether Lakatos might not have gone on to develop or endorse AOE if he had lived.

Chapter Five

The Need for a Revolution in the Philosophy of Science

(First published in the *Journal for General Philosophy of Science* vol. 33, 2002, 381-408.)

Abstract

There is a need to bring about a revolution in the philosophy of science, interpreted to be both the academic discipline, and the official view of the aims and methods of science upheld by the scientific community. At present both are dominated by the view that in science theories are chosen on the basis of empirical considerations alone, nothing being permanently accepted as a part of scientific knowledge independently of evidence. Biasing choice of theory in the direction of simplicity, unity or explanatory power does not permanently commit science to the thesis that nature is simple or unified. This current "paradigm" is, I argue, untenable. We need a new paradigm, which acknowledges that science makes a hierarchy of metaphysical assumptions concerning the comprehensibility and knowability of the universe, theories being chosen partly on the basis of compatibility with these assumptions. Eleven arguments are given for favouring this new "paradigm" over the current one.

1 Introduction

There is an urgent need for a revolution in the philosophy of science. By "the philosophy of science" I mean not only the academic discipline, but also the view about what the aims and methods of science are (and ought to be) upheld by the scientific community, influencing the way science is pursued, taught, funded, discussed, communicated to the public. I mean the view that is embedded, as it were, in the actual institutional structure of science, shaping the way science proceeds, as well as the academic discipline (at present linked only tenuously to science itself, and somewhat despised by many working scientists).

Philosophers of science are of course aware of many diverse "philosophies of science", views about what the aims and methods of science are, or ought to be. Logical positivism, inductivism, logical empiricism, hypothetico-deductivism, falsificationism, conventionalism, constructive empiricism, pragmatism, realism, induction-to-the-best-explanationism, the views of Kuhn, Lakatos and Feyerabend are just some of the available options (to say nothing of the views of sociologists and historians of science). However, these views, diverse as they may seem to be, have one basic thesis in common. I shall call this thesis *standard empiricism*.

2 The Old Paradigm: Standard Empiricism

Standard Empiricism (SE) asserts: In science, ideally, all claims to knowledge are to be assessed impartially with respect to the evidence, the simplicity, unity or explanatory power of theories being taken into account as well, *no thesis about the world being upheld permanently as a part of knowledge independently of evidence, let alone in violation of evidence*. Most, if not all, versions of SE stress that questions of simplicity, unity and explanatory power play a valid, important role in influencing choice of theory in science, in addition to considerations of empirical success – although some versions of SE give to simplicity and explanatory power much more important roles in science than other versions do. The decisive point that all versions of SE agree on is that no substantial thesis about the nature of the universe can be upheld as a part of scientific knowledge

independently of empirical considerations, and certainly not in violation of empirical considerations. In so far as theory choice is biased in the direction of simplicity, unity or explanatoriness, this bias must not commit science to making the permanent assumption that nature herself is simple, unified or explainable.

This rather thin thesis is common ground for logical positivism, inductivism, logical empiricism, hypothetico-deductivism, falsificationism, conventionalism, constructive empiricism, pragmatism, realism, induction-to-the-best-explanationism, and the views of Kuhn and Lakatos.¹¹⁰ There is a sense in which even Feyerabend, and even social constructivist and relativist sociologists and historians of science uphold SE as the best available ideal of scientific rationality. *If* science can be exhibited as rational, they hold (in effect), then this must be done in a way that is compatible with SE. The failure of science to live up to the rational ideal of SE is taken by them to demonstrate that science is not rational. That it is so taken demonstrates convincingly that they hold SE to be the only possible rational ideal for science (an ideal which cannot, it so happens, in their view, be met).

SE is more or less unthinkingly taken for granted by the vast majority of working scientists – so much so that it is rather rare to find the doctrine being explicitly formulated, let alone defended. Scattered throughout the writings of scientists one can, nevertheless, find affirmations of the view. Thus Planck once remarked "Experiments are the *only* means of knowledge at our disposal. The rest is poetry, imagination" (Atkins, 1983, p. xiv). Or, as Poincaré (1952, p. 140) put it "Experiment is the sole source of truth. It alone can teach us something new; it alone can give us certainty."¹¹¹

¹¹⁰. For discussion of the claim that Kuhn and Lakatos defend versions of SE see Maxwell (1998), p. 40. Bayesianism might seem to reject SE, in acknowledging both prior and posteriori probabilities. But Bayesianism tries to conform to the spirit of SE as much as possible, by regarding prior probabilities as personal, subjective and non-rational, their role in theory choice being reduced as rapidly as possible by empirical testing: see Maxwell (1998), p. 44.

¹¹¹. For more detailed discussion of the point that SE is widely taken for granted see Maxwell (1984), chs. 2 and 6, and Maxwell (1998), ch. 2.

Despite all this, SE can easily be shown to be untenable, in a quite decisive fashion. In the next section I spell out reasons for rejecting the prevailing "paradigm" of SE. In the section after I expound the new "paradigm" which, I hold, ought to replace SE. After some pertinent remarks about what it means to assert that the universe is "physically comprehensible", or has a "unified dynamic structure", I then give eleven reasons for rejecting SE and accepting in its stead the new paradigm of *aim-oriented empiricism*.

All this restates and, in some respects elaborates on arguments to be found in Maxwell (1998).¹¹²

3 Refutation of Standard Empiricism

Given any scientific theory, however well verified empirically, there will always be infinitely many rival theories which fit the available evidence just as well, but which make different predictions, in an arbitrary way, for phenomena not yet observed. Thus, given Newtonian theory (NT), one rival theory might assert: everything occurs as NT asserts up till midnight tonight when, abruptly, an inverse cube law of gravitation comes into operation. A second rival theory might assert: everything occurs as NT asserts, except for the case of any two solid gold spheres, each having a mass of a thousand tons, moving in otherwise empty space up to a mile apart, in which case the spheres attract each other by means of an inverse cube law of gravitation. A third rival asserts that everything occurs as NT asserts until three kilograms of gold dust and three kilograms of diamond dust are heated in a platinum flask to a temperature of 450°C, in which case gravitation will instantly become a repulsive force everywhere. And so on. There is no limit to the number of rivals to NT that can be concocted in this way, each of which has all the predictive success of NT as far as observed phenomena are concerned but which makes different predictions for some as yet unobserved phenomena.¹¹³ Such theories can even be concocted which are

¹¹². See also Maxwell (1972a), (1974), (1984) and (1993).

¹¹³. All the possible phenomena, predicted by any dynamical physical theory, T, may be represented by an imaginary "space", S, each point in S corresponding to

more empirically successful than NT, by arbitrarily modifying NT, in just this entirely *ad hoc* fashion, so that the theories yield correct predictions where NT does not, as in the case of the orbit of Mercury for example (which very slightly conflicts with NT).¹¹⁴

One can set out to refute these rival theories by making the relevant observations or experiments, but this needs an infinitely long time to complete as there are infinitely many rival theories to be refuted, each requiring a different refuting experiment. Thus, if science really did take seriously the idea that evidence alone decides what theories are to be accepted and rejected, scientific knowledge would be drowned in an infinite ocean of rival theories, all just as empirically successful as currently accepted theories, or actually even more successful empirically. Science would come to an end.¹¹⁵

Why does this not happen in scientific practice? Because, as most versions of SE stress, in practice *two* considerations govern acceptance and rejection of theories in science: (1) considerations of empirical success and failure; and (2) considerations that have to do with the simplicity, unity or explanatory power of the theories in question. In order to be accepted as a part of scientific knowledge, a theory must satisfy *both* considerations. It must be

a particular phenomenon, a particular kind of physical system evolving in time in the way predicted by T. In order to specify *ad hoc* rivals to T that fit all available evidence just as well as T does, all we need do is specify a region in S that consists of phenomena that have not been observed, and then replace the phenomenon predicted by T with anything we care to think of. Given any T, there will always be infinitely many such *ad hoc* rivals to T.

¹¹⁴. For a more detailed discussion of empirically successful *ad hoc* rivals to accepted theories, see Maxwell (1998), pp. 51-54.

¹¹⁵. This argument generalizes Goodman's (1954) argument concerning green and grue. Two rival theories considered by Goodman are "All emeralds are green" and "All emeralds are grue", where an emerald is grue if it is examined before time *t* and green or not examined before *t* and blue. Before time *t*, available evidence appears to support both theories equally well. The argument given here improves on Goodman's argument, in my view, in that it makes closer contact with science. *Ad hoc* theories, admittedly not quite as bizarre as the rivals to NT that I have indicated, can be a serious issue in science. It is important to appreciate that the problem of why such theories deserve to be rejected is a serious problem for science, and not merely a weird philosophical puzzle.

both empirically successful *and* simple, unified, or explanatory in character.¹¹⁶

Scientific theories that are accepted as a part of scientific knowledge, NT let us say, classical electromagnetism, quantum theory or Einstein's theories of special and general relativity, do (more or less adequately) satisfy *both* considerations. They are both amazingly successful in their capacity to predict observable phenomena, and astonishingly simple, unified, explanatory.

But the infinitely many empirically successful rivals to these accepted theories all *fail* to satisfy the second consideration. They may fit all available evidence just as well as Newton's theory does, or Einstein's theories do: but they fail, quite drastically, to be simple, unified, explanatory. For these rival theories all assert that, for some as yet unobserved kind of phenomenon, something entirely peculiar and arbitrary occurs. Where NT assures us that gravitation obeys an inverse square law and is attractive uniformly everywhere, for all time, the aberrant rivals to NT assert that for some specific kind of phenomenon or range of phenomena gravitation obeys a quite different law, an inverse cube law perhaps, or one that asserts that gravitation is a repulsive rather than attractive force.

Thus the infinitely many rivals to accepted physical theories are rejected out of hand, not on empirical grounds, but because they are grotesquely *ad hoc*, grotesquely lacking in simplicity, unity, explanatory power.

This, then, is why in practice science is not buried beneath an infinite mountain of rival theories, all of which fit all available evidence equally well, if not better. Almost all the rivals are horribly complex, disunified, non-explanatory.

But now comes the decisive point. In persistently rejecting infinitely many such empirically successful but grotesquely *ad hoc* theories, science in effect makes a big permanent assumption about the nature of the universe, to the effect that it is such that no grotesquely *ad hoc* theory is true, however empirically successful

¹¹⁶. Induction-to-the-best-explanation gets this part right!

it may appear to be for a time.¹¹⁷ Without some such big assumption as this, the empirical method of science collapses. Science is drowned in an infinite ocean of empirically successful *ad hoc* theories.

The orthodox conception of science is, in short, untenable.¹¹⁸

4 The New Paradigm: Aim-Oriented Empiricism

At once the question arises: Granted that science must make some kind of big assumption about the nature of the universe if it is to be possible at all, what precisely ought this assumption to be, and on what basis is it to be made? If science is to proceed successfully we must make an assumption that is near enough correct; and yet it is just here that we are most ignorant, and are almost bound to get things hopelessly wrong.

The solution to this fundamental epistemological problem of science (the very existence of which is denied by SE) is to construe science as adopting, as a part of scientific knowledge, a hierarchy of cosmological assumptions about the comprehensibility and knowability of the universe (as we saw in chapter 4), these assumptions asserting less and less about the universe as one ascends the hierarchy, thus being more and more likely to be true: see diagram 2. Corresponding to these cosmological assumptions there are methodological rules (represented by dotted lines in diagram 2) which govern acceptance of assumptions lower down in the hierarchy, and which, together with empirical considerations, govern acceptance and rejection of scientific theories. The top assumption, at level 7, is such that accepting this assumption as a part of scientific knowledge can only aid, and can never damage science (or the task of acquiring knowledge more generally) whatever the universe may be like. This is justifiably a permanent item of scientific knowledge. Thus at level 7 we have the thesis that the universe is such that we can acquire some knowledge of our local circumstances: we are justified in accepting

¹¹⁷. This is where "induction-to-the-best-explanation" goes wrong. It tries to make persistent preference for explanatory theories in science, independent of empirical considerations, something that is compatible with SE.

¹¹⁸. For a much more detailed presentation of this refutation of SE see Maxwell (1998), ch. 2.

this as a permanent part of scientific knowledge. As we descend, from level 7 to level 3, the corresponding theses make increasingly substantial assertions about the nature of the universe: it becomes increasingly likely that these theses are false. At each level, from 6 to 3, we adopt that assumption which (a) is compatible with the assumption above it in the hierarchy (in so far as this is possible), and (b) holds out the greatest hope for the growth of empirical knowledge, and seems best to support the growth of such knowledge (at levels 1 and 2). If currently adopted cosmological assumptions, and associated methods, fail to support the growth of empirical knowledge, or fail to do so as apparently successfully as rival assumptions and methods, then assumptions and associated methods are changed, at whatever level appears to be required.¹¹⁹ In this way we give ourselves the best hope of making progress, of acquiring authentic knowledge, while at the same time minimizing the chances of being taken up the garden path, or being stuck in a cul de sac. The hope is that as we increase our knowledge about the world we improve the cosmological assumptions implicit in

¹¹⁹. It may be asked: But how can acceptance of a level 3 assumption both influence, and be influenced by, acceptance of level 2 theories? The answer is that, at any stage in the development of science, rival level 3 ideas can contend; these lead to rival research programmes (Lakatos, 1970), which can be assessed with respect to their relative empirical growth. Within a research programme, theories are rejected that clash with the basic level 3 idea; this idea is rejected if a rival research programme meets with greater empirical success over a period of time. Level 3 ideas are also assessed in terms of how well they exemplify the accepted level 4 thesis. (But this too is open to revision, if such a revision leads to a more empirically progressive research programme.)

Suppose that B_1 is accepted at level 3, at a certain stage in the development of science. Suppose, now, that successive, increasingly empirically successful (level 2) theories, T_1, T_2, T_3 are put forward which clash with B_1 , but exemplify beautifully a different level 3 thesis B_2 , so that the B_2 research programme achieves greater empirical success than the B_1 programme does. Suppose further that B_1 and B_2 exemplify the level 4 thesis (of physical comprehensibility) equally well. This justifies accepting B_2 and rejecting B_1 (on quasi-empirical grounds). If B_2 exemplifies the level 4 thesis (of physical comprehensibility) better than B_1 does, all the more reason to accept B_2 in preference to B_1 . For further details of how metaphysical theses are to be selected, at various levels, partly on the basis of the empirical success and failure of rival research programmes, within the framework of aim-oriented empiricism, see Maxwell (1998), chs. 4 and 5.

our methods, and thus in turn improve our methods. As a result of improving our knowledge we improve our knowledge about how to improve knowledge. Science adapts its own nature to what it learns about the nature of the universe, thus increasing its capacity to make progress in knowledge about the world – the methodological key to the astonishing, accelerating progress of modern science.

This conception of science, postulating more or less specific evolving aims and methods for science within a framework of more general fixed aims and methods, I call *aim-oriented empiricism*.¹²⁰ It is a special case of a more general idea of *aim-oriented rationality*, according to which, whenever basic aims are problematic (as they usually are in science and in life) we need to display aims at distinct levels of specificity and generality, thus creating a framework within which we have the best chance of improving more or less specific, problematic aims-and-methods as we proceed, in the light of success and failure.¹²¹

According to aim-oriented empiricism (AOE), then, scientific knowledge can be represented (in a highly schematic and simplifying way) as being made up of the following seven levels: see diagram 2. At level 1, we have empirical data (low level experimental laws). At level 2, we have our best fundamental physical theories, currently general relativity and the so-called standard model. At level 3, we have the best, currently available

¹²⁰. Corresponding to each cosmological thesis, at level 3 to 10, there is a more or less problematic *aim* for theoretical physics: to specify that cosmological thesis as a true, precise, testable, experimentally confirmed "theory of everything". Aims corresponding to levels 9 and 10 are relatively unproblematic: circumstances will never arise such that it would serve the interests of acquiring knowledge to revise these aims. As one descends the hierarchy of cosmological assumptions, the corresponding aims become increasingly problematic, increasingly likely to be unrealizable, just because the corresponding assumption becomes increasingly likely to be false. Whereas upper level aims and methods will not need revision, lower level aims and methods, especially those corresponding to level 3, will need to be revised as science advances. Thus lower level aims and methods evolve within the fixed framework of upper aims and methods.

¹²¹. For the generalization of aim-oriented empiricism to form aim-oriented rationality see Maxwell (1984 or 2007), (1992), (2000a), (2002a).

specific idea as to how the universe is physically comprehensible. This asserts that everything is made of some specific kind of physical entity: corpuscle, point-particle, classical field, quantum field, convoluted space-time, string, or whatever. Because the thesis at this level is so specific, it is almost bound to be false (even if the universe is physically comprehensible in some way or other). Here, ideas evolve with evolving knowledge. At level 4 we have the much less specific thesis that the universe is *physically* comprehensible in some way or other (a thesis I shall call *physicalism*¹²²); and at level 5 we have the even less specific thesis that the universe is *comprehensible* in some way or other, whether physically or in some other way. At level 6 there is the even less substantial thesis that there is some rationally discoverable thesis about the nature of the universe which, if accepted, makes it possible progressively to improve methods for the improvement of knowledge ("rationally discoverable" meaning at least that the thesis is not an arbitrary choice from infinitely many analogous theses). At level 7 there is the thesis that the universe is such that we can acquire knowledge of our local circumstances sufficient to enable us to continue to live. This is a thesis we can accept whatever the universe may be like, since accepting it as a part of knowledge can never harm the pursuit of truth. If it is false, we have had it whatever we assume.¹²³

Ideally, the thesis at level 2 implies the one at level 3, and so on up the hierarchy until one reaches level 7. This is true for levels 4 to 7. It breaks down dramatically, however, when we come to levels 2, 3 and 4. Fundamental theories currently accepted in physics, general relativity and the standard model, clash, and thus fail to exemplify physicalism. Furthermore, instead of postulating just one kind of self-interacting entity, the standard model postulates three kinds of forces, and many different kinds of

¹²². Smart (1963) has used the term 'physicalism' to stand for the view that the world is made up entirely of physical entities of the kind postulated by fundamental physical theories – electrons, quarks and so on. As I am using the term, 'physicalism' stands for the much stronger doctrine that the universe is physically comprehensible, that it is such that some yet-to-be-discovered, unified "theory of everything" is true.

¹²³. For further details see Maxwell (1998, ch. 1; 2007, ch. 14).

particles with diverse properties, such as mass, that are not theoretically determined. All this is a sign of our ignorance (just as failure of theories to predict phenomena successfully is). What drives physics forward is the attempt to solve the problems that arise as a result of clashes between levels 1, 2, 3 and 4. According to AOE, a basic task of theoretical physics will have been completed when a level 2 theory has been discovered which (a) in principle predicts all (physically) possible level 1 phenomena, and (b) implies a true level 3 thesis, which (c) exemplifies (and thus implies) the level 4 thesis of physical comprehensibility (physicalism).

5 Physical Comprehensibility

Two key theses in the hierarchy just indicated are the level 5 thesis that the universe is comprehensible, and the level 4 thesis that the universe is *physically* comprehensible. What do these assert?

The level 5 thesis of comprehensibility asserts that the universe is such that there is *something* (God, tribe of gods, cosmic goal, pattern of physical law, cosmic programme or whatever), which exists everywhere in an unchanging form and which, in some sense, determines or is responsible for everything that changes (all change and diversity in the world in principle being explicable and understandable in terms of the underlying unchanging *something*). A universe of this type deserves to be called "comprehensible" because it is such that everything that occurs, all change and diversity, can in principle be explained and understood as being the outcome of the operations of the one underlying *something*, present throughout all phenomena.

If the *something* that determines all change is what corresponds out there in the world to a unified pattern of physical law, then the universe is physically comprehensible. The universe is physically comprehensible, in other words, if some yet-to-be-discovered unified physical "theory of everything" or "final theory" is true.

Let us now consider, in a little more detail, what this means. (It is important to appreciate that, in what follows, I am indicating, informally, what it is that a specific *thesis* – physicalism – asserts about the universe; I am not so much analyzing the meaning of

"unity", as explicating what it is that physicalism asserts in asserting: the universe has a unified dynamic structure: for further details, see Maxwell, 1998, chs. 3 and 4.)

The basic idea is that the physically comprehensible universe consists of two aspects, U and V. U is an aspect of the cosmos that is present throughout all phenomena in an unchanging form, while V is that aspect that varies and changes from place to place and time to time. For physical comprehensibility we require first, that U is unified, and second, that U determines (perhaps probabilistically) the way in which V changes. But what does "unity" mean here? And what does it mean to assert that this unified aspect "determines" change?

As a first stab at answering these questions, let us consider an elementary example of a universe that is physically comprehensible in the specified sense. Consider a universe that consists only of the classical electromagnetic field in empty space (there being no charged particles to create, or be acted on, by the field).¹²⁴ This field is a physical entity that is spread out smoothly throughout space and time. It consists of two inter-related fields, the electric and magnetic fields. These inter-related fields together have two aspects, U and V.

The aspect that varies, V, (to take them in reverse order) consists of varying strengths or intensities of the electric field, **E**, and magnetic field, **B**.

The way in which **E** and **B** change is precisely determined by James Clerk Maxwell's equations of the electromagnetic field. That is, given the state of the field, the values of **E** and **B** at each point, at some instant (in some reference frame), this together with Maxwell's equations determines uniquely all subsequent (and previous) states of the field.

Maxwell's equations for the electromagnetic field in the vacuum assert:

$$(1) \nabla \cdot \mathbf{E} = 0 \qquad (3) \nabla \times \mathbf{E} = - \frac{\partial \mathbf{B}}{\partial t}$$

¹²⁴. This toy example of a physically comprehensible universe does not exhibit full dynamic unity, in the sense indicated below, in that the field and space-time are not unified to form *one* entity.

$$(2) \nabla \cdot \mathbf{B} = 0 \qquad (4) \nabla \times \mathbf{B} = - \frac{\partial \mathbf{E}}{c \partial t}$$

where c is the velocity of light.

Ordinarily (1) to (4) would be thought of as factual assertions. Here, they are to be reinterpreted as *analytic* statements, which tell us what it is for something to be the electromagnetic field. The empirical import of the theory is contained in a fifth postulate:

(5) The classical electromagnetic field, as characterized by (1) to (4), exists everywhere (and this is all that exists).

(1) to (4) are to be regarded as specifying a dispositional or necessitating physical property: if it exists everywhere then, necessarily, that which changes, values of the electric and magnetic fields, change in accordance with (1) to (4). *If* this necessitating property exists everywhere at an instant, then it, together with the instantaneous values of the electric and magnetic fields everywhere, determine all subsequent states of the field. In this sense the necessitating property (of being the classical electromagnetic field), specified by (1) to (4), itself does not change but determines how varying values of \mathbf{E} and \mathbf{B} do change.

Any fundamental physical theory that specifies how postulated fundamental physical entities evolve and interact can be interpreted in this "essentialistic" way. Instead of basic laws being interpreted as specifying regularities observed by the entities, they can be reinterpreted as specifying necessitating *properties* that the entities possess which determine (necessarily) how the entities evolve and interact. To say that the universe is physically comprehensible is to say that there exists an unchanging *something* (capable of being specified by a unified, essentialistically interpreted physical theory) which exists everywhere and which determines (perhaps probabilistically) the way in which everything that changes does change.¹²⁵

¹²⁵. For further details concerning this essentialistic approach to physics, and the hypothetical existence of necessary connections between successive events, see: Maxwell: (1968); (1993), Part 2, pp. 81-101; (1998), pp. 141-155.

What, it may be asked, leads one to hold that there is here one unified entity, the electromagnetic field, and not two entities, the electric field and the magnetic field? There are three points to note.

First, (3) and (4) specify how a magnetic field changing with respect to time creates a circulating electric field, and an electric field changing with respect to time creates a circulating magnetic field. The two fields, \mathbf{E} and \mathbf{B} , are not distinct; a change in the one creates the other. Interestingly enough, it is this interdependence of the electric and magnetic fields that is responsible for the other great feat of unification achieved by classical electrodynamics, namely the unification of light and electromagnetism. What these two equations imply is that once periodic changes in the electric or magnetic field have been set up they travel through space with the velocity of light, changing \mathbf{E} producing a changing \mathbf{B} , in turn producing a changing \mathbf{E} , and so on. This, according to classical electrodynamics, is what light is.

The second point to note is the symmetry that exists between the way \mathbf{E} affects \mathbf{B} , and \mathbf{B} affects \mathbf{E} .

And the third point is that, given any specific space-time chunk of electromagnetic field evolving in accordance with the above four postulates, the way in which the field divides up into \mathbf{E} and \mathbf{B} depends on the choice of reference frame. The lengths and directions of vectors representing \mathbf{E} and \mathbf{B} , for the same chunk of space-time, differ with respect to two reference frames travelling at some uniform velocity with respect to each other. But, according to special relativity, nothing of absolute (or theoretically fundamental) significance can depend on the choice of reference frame. Any specific way of dividing up the electromagnetic field into \mathbf{E} and \mathbf{B} has as much absolute significance as a choice of velocity for some object. We cannot, in short, think of the electromagnetic field as being made up of two distinct fields, \mathbf{E} and \mathbf{B} , since any specific choice of \mathbf{E} and \mathbf{B} is arbitrary in that it amounts to an arbitrary choice of reference frame. There is thus the *one* entity, the electromagnetic field, made up of two symmetrically interdependent aspects, \mathbf{E} and \mathbf{B} .

There are two rather different kinds of unification in theoretical physics, which may be called "unification by annihilation" and

"unification by synthesis". Unification by annihilation occurs when it emerges that, instead of there being a number of distinct kinds of entity (substance or phenomenon), E_0, E_1, \dots, E_N , there is just the one kind of entity, E_0 , the other entities, E_1, \dots, E_N being no more than aspects of E_0 . Unification by synthesis occurs when it emerges that apparently distinct kinds of entities, E_1, \dots, E_N are really different aspects or facets of one common entity, $E_1 + \dots + E_N$.

Classical electrodynamics illustrates both kinds of unification. On the one hand (as I have just mentioned) classical electrodynamics depicts light as nothing more than wave-like changes in the electromagnetic field: light, as a phenomenon distinct from the electromagnetic field, is annihilated. On the other hand, the theory reveals that the electric and magnetic forces, which before the advent of the theory seemed to be two distinct forces, are two facets of one entity, the electromagnetic force or field: the electric and magnetic forces are unified by synthesis.

Great steps of unification by annihilation from the history of theoretical physics include: the discovery that the millions of different substances that exist are made up of different chemical combinations of under one hundred different elements; the discovery that these distinct elements are made up of atoms in turn made up of just three kinds of particle: the electron, proton and neutron; the discovery that gravitation is nothing more than the curvature of space-time; the discovery that the dozens of different hadronic particles revealed in the 1950's and 1960's are nothing more than relatively few different kinds of quarks interacting by means of the gluons associated with the strong force. Great steps of unification by synthesis include: the partial unification of space and time achieved by Einstein's special theory of relativity; the partial unification of energy and matter achieved, again, by special relativity, and enshrined in the most famous equation of modern physics: $E = mc^2$; the partial unification of the electromagnetic and so-called weak force achieved by Weinberg's and Salam's theory of quantum electroweak theory; the unification of the eight gluons of the strong force achieved by quantum chromodynamics.¹²⁶

¹²⁶. For a slightly more detailed, informal account of these great feats of unification see Maxwell (1998), pp. 123-140.

To sum up: the thesis that the universe is physically comprehensible asserts that it is made up of two aspects, U, a necessitating property which exists everywhere and does not change, and V, which is the aspect of the universe that is different, from place to place and time to time. Given U, and given V at an instant throughout the universe, all subsequent states of V are precisely determined (perhaps probabilistically). Furthermore, U is unified in the sense that, in so far as it consists of distinct parts or aspects, U_1, U_2, \dots, U_N , these are unified by means of a symmetry in a way that is analogous to how the electric and magnetic fields are unified by the symmetry of special relativity (Lorentz invariance), in the manner indicated above.¹²⁷

6 Eleven Reasons for Rejecting Standard Empiricism and Adopting Aim-Oriented Empiricism Instead

Arguments have already been given for rejecting SE and accepting AOE in its stead. Here are eleven further arguments.¹²⁸

(1) *Greater rigour*. An elementary requirement for intellectual rigour is that assumptions that are substantial, influential, problematic and implicit need to be made explicit, so that they can be criticized and, we may hope, improved. AOE puts this into practice, in that metaphysical assumptions, implicit in persistent

¹²⁷. Additional examples of the way unity is achieved by symmetry are: the (partial) unity of the electroweak theory, due to the fact that the theory exhibits the local gauge symmetry of $U(1) \times SU(2)$; and the unity of chromodynamics due to the fact that the theory exhibits the local gauge symmetry of $SU(3)$. For full unity, we require that the symmetry group is not a direct product of subgroups. The symmetry group of the electroweak theory *is* a direct product of subgroups; hence the theory does not fully unify the electromagnetic and weak forces. It may be that the symmetry is spontaneously broken, as in the case of electroweak theory. Within modern theoretical physics, the hunt for the precise nature of U is largely (but not entirely) the hunt for the symmetry group that bestows unity on U. For an informal exposition of these matters for non-experts, and for further literature on the subject, and for a more detailed discussion of the meaningfulness of physicalism, see Maxwell (1998), especially chs. 3 and 4, and the appendix.

19. In what follows, there is some overlap with points made in the previous chapter. There is, however, a big difference. In the previous chapter, AOE is compared with Popper's falsificationism; in this chapter, the comparison is with SE, which includes many views in addition to falsificationism.

preference in science for simple, unified or explanatory theories, are made explicit in such a way that they can be subjected to maximum critical appraisal, AOE providing a framework for their critical assessment. SE denies that any such metaphysical assumptions are implicit in the persistent scientific preference for simple theories. The attempt to do science in accordance with SE thus undermines scientific rigour, in that it leads science to violate the above elementary requirement for rigour. This greater rigour of AOE is no mere formal matter: it makes it possible for there to be a rational method of discovery in theoretical physics, something which SE does not permit, as we shall see (point (9) below).

(2) *Solution to Fundamental Epistemological Problem of Science.*

Once it is recognized that persistent preference for simple or unifying theories in physics independent of empirical considerations, or even against such considerations, means that science must make a big, permanent assumption about the nature of the universe, the fundamental question arises: What is this permanent assumption, and on what basis is it made? SE not only fails to solve the problem; it denies the very existence of the problem by denying that any such assumption *is* made by science in persistently preferring simple or unifying theories. AOE, by contrast, acknowledges the problem, stresses that this is the fundamental epistemological problem of science, and solves it by postulating a hierarchy of increasingly contentless assumptions concerning the comprehensibility and knowability of the universe, the top two being such that it can only benefit, and cannot harm, the search for knowledge whatever the universe is like. At any level in the hierarchy below the top two levels, that assumption is selected which (a) best accords with the assumption above, and (b) is the most fruitful, or promises to be the most fruitful, from the standpoint of the growth of empirical knowledge, at levels 1 and 2 (see diagram). Corresponding to these cosmological assumptions there are methodological rules (represented by dotted lines in diagram) which govern acceptance of assumptions lower down in the hierarchy, and which, together with empirical considerations, govern acceptance and rejection of scientific theories. This hierarchical structure of assumptions and associated methods is put forward as giving science the best hope of discovering and

accepting those assumptions and methods which best promote the growth of empirical knowledge, at the same time giving science the best hope of discovering and rejecting assumptions that are blocking such growth. The hierarchical framework enables science to adapt its aims and methods to what it finds out about the universe, improving knowledge leading to improvement in knowledge about how to improve knowledge. All this is suppressed by SE.

(3) *The nature of simplicity.* SE cannot solve the problem of what the simplicity, unity or explanatory character of a theory *is*, a problem that even Einstein found baffling.¹²⁹ This problem arises because a simple theory can always be reformulated so that it becomes complex, and *vice versa*.¹³⁰ AOE solves the problem without difficulty. The totality of fundamental physical theory, T, is unified to the extent that its *content* exemplifies physicalism. The more the content of T departs from exemplifying physicalism, the more disunified T is.¹³¹ Because what matters is content, not form, the way T is *formulated* is irrelevant to this way of assessing simplicity or unity.¹³² SE cannot avail itself of this way of

^{129.} See Einstein (1969), p. 23. For discussion, see Maxwell (1998), pp. 105-106.

^{130.} For accounts of failed attempts at solving the problem of simplicity within the framework of SE, see Salmon (1989), Maxwell (1998), pp. 56-68.

^{131.} Dynamical theories are partially ordered with respect to the extent that they exemplify physicalism, with respect to their degree of unity, in other words. For further details see Maxwell (1998), ch. 4.

^{132.} It may be objected that this proposed solution to the problem of simplicity is circular: the unity of level 2 theory is explicated in terms of the unity of level 4 physicalism. But this objection is not valid. In order to solve the problem, it is not necessary to explicate what "simplicity" or "unity" mean; rather, what needs to be done is to show how theories can be partially ordered with respect to "simplicity" or "unity" in a way that does not depend on formulation. This is achieved by partially ordering theories in terms of how well their content exemplifies the content of physicalism, so that, roughly, the more the content of a theory violates the symmetries associated with the content of physicalism, the less unity it has. As long as physicalism is a meaningful thesis, and provides a formulation-independent way of partially ordering theories in the way indicated, this suffices to solve the problem. That physicalism embodies intuitive ideas of "unity" is a bonus. For a more detailed rebuttal of this objection, see Maxwell (1998), pp. 118-123.

assessing unity because it involves acknowledging that physicalism is a basic tenet of scientific knowledge, something which SE denies. Within AOE, there is a second way in which the unity of T may be assessed: in terms of the extent to which the content of T exemplifies the best available more or less specific idea as to how the universe is physically comprehensible, at level 3. This second conception of simplicity or unity evolves with the evolution of level 3 ideas. As we improve our ideas about how the universe is unified, with the advance of knowledge in theoretical physics, so non-empirical methods for selecting theories on the basis of simplicity or unity improve as well. Thus current symmetry principles of modern physics, such as Lorentz invariance and gauge invariance, which guide acceptance of theory, are an advance over simplicity criteria upheld by Newton. This account of simplicity can be extended to individual theories in two ways. First, we may treat an individual theory as a candidate theory of everything. Second, given two individual theories, T_1 and T_2 , and given the rest of fundamental theory, T, T_1 is simpler than T_2 iff $T + T_1$ is simpler than $T + T_2$, where the latter is assessed in one or other of the ways indicated above.¹³³

(4) *The nature of scientific method.* Most versions of SE acknowledge (correctly) that two considerations govern selection of theory in science, namely considerations that have to do with (a) evidence, and (b) simplicity. But because SE cannot solve the problem of what simplicity *is*, SE cannot, with any precision, specify what methods are involved when theories are selected on the basis of simplicity. Nor can SE do justice to the way in which the methods of physics evolve with evolving knowledge, especially methods that assert that acceptable theories must satisfy this or that *symmetry*. AOE, on the other hand, solves the problem of simplicity, and thus can specify precisely what methods are involved when theories are selected on the basis of simplicity. Furthermore, AOE can do justice to *evolving* criteria of simplicity (as we have just seen). According to AOE, the totality of

¹³³. For a very much more detailed exposition of this solution of the problem of simplicity, together with an account of the way in which great unifying theories of physics illustrate the solution, see Maxwell (1998), chs. 3 and 4.

fundamental physical theory, T, can be assessed with respect to how well its content exemplifies (i) the relatively fixed level 4 thesis of physicalism, or (ii) the evolving, best available level 3 thesis. Whereas (i) constitute fixed criteria of simplicity or unity (as long as physicalism is not modified), (ii) constitute evolving criteria, criteria of unity that improve with improving knowledge.

(5) *Justification of the methods of physics.* How are the methods, M, of science to be *justified*? That is, what justifies the claim that theories, T, accepted in accordance with methods, M, deserve to be regarded as constituting knowledge, in some sufficiently good sense of "knowledge"?

Attempts to solve this problem of justification, in essence the problem of induction, within the framework of SE, have all failed. This is not surprising: we have just seen that SE cannot even solve the preliminary problem of specifying what the methods, M, of science *are*. If M cannot be specified adequately, there is small hope that they can be justified. By contrast, AOE reveals that science provides its own justification for its choice of methods, as it proceeds.

Assumptions at the top two levels of the hierarchy are accepted because this can only help, and cannot harm, the search for knowledge whatever the universe may be like. Assumptions lower down in the hierarchy are accepted because these assumptions seem to promote the growth of empirical knowledge, at levels 1 and 2, better than any rival assumptions. This kind of ostensible fruitfulness for the (ostensible) growth of empirical knowledge, is the best indication we can have that the assumptions we are making are along the right lines. Displaying these assumptions in the form of a hierarchy, each assumption having decreasing content as we go up the hierarchy, gives science the best hope of discovering precisely where in the hierarchy an assumption is false, and needs to be modified or replaced.

Any claim to factual knowledge, however trivial, has a cosmological dimension. In claiming to know that I can walk across a room, I thereby claim that the entire cosmos is such that no vast explosion is occurring anywhere which will spread with near infinite speed to engulf the room before I have crossed it. The hierarchy of cosmological assumptions postulated and specified by

AOE makes explicit for science a cosmological dimension that is implicit in all claims to factual knowledge.

The solution to the problem of induction must justify scientific knowledge *to the extent that this is possible*. AOE does just that. AOE does not, perhaps, enormously increase confidence in the correctness of science: but it does considerably increase our understanding of science. Some may reject the claim that AOE solves the problem of induction on the grounds that AOE does fail to increase confidence in science. But AOE cannot justifiably be found wanting for failing to provide what cannot be had. It is more rational to acknowledge the real fallibility of scientific knowledge, than to delude oneself into believing scientific knowledge is more secure than it is.¹³⁴ (Only those who suffer from this delusion will be unconvinced by the claim that AOE solves the problem of induction.)

(6) *Justification of preference for simple theories*. Not only does SE fail to say what simplicity is; it also fails to justify persistent preference for simple theories in science. AOE solves both problems. Persistent preference for simple, unifying or explanatory theories is justified because these are the theories which best exemplify the best available assumptions at levels 3 and 4. Science is justified in accepting these assumptions for reasons indicated in (5).

(7) *Evolving aims-and-methods*. A striking feature of physics is the way non-empirical methods, determining what theories will be accepted and rejected, have evolved from Newton's time to today. Newton, in his *Principia*, formulated four rules of reasoning, three of which are concerned with simplicity. Principles that have been proposed since his day include: invariance with respect to position, orientation, time, uniform velocity, charge conjugation, parity,

¹³⁴. How secure one judges theoretical scientific knowledge to be ought to depend on how secure one judges the theses of comprehensibility, and physical comprehensibility, to be. If these theses are held to be no more than wild metaphysical speculations, theoretical scientific knowledge ought to be held to be similarly speculative. But if it is legitimate to hold that these theses are reasonably secure items of scientific and common sense knowledge, then theoretical scientific knowledge may legitimately be held to be reasonably secure also.

time-reversal; principles of conservation of mass, momentum, angular momentum, energy, charge; Lorentz invariance; Mach's principle, the principle of equivalence; principles of gauge invariance, global and local; supersymmetry; duality principles; the principle that different kinds of particle should be reduced to one kind, and different kinds of force should be reduced to one kind; the principle that space-time on the one hand, and particles-and-forces on the other, should be unified. All of these principles can be interpreted as methodological rules which specify requirements theories must meet in order to be accepted. They can also be interpreted as physical principles, making substantial assertions about such things as space, time, matter, force. Some, such as conservation of mass, parity, and charge conjugation, have been shown to be false; others, such as Mach's principle, have never been generally accepted; still others, such as supersymmetry, remain speculative.

Principles such as these, which can be interpreted either as physical assertions or as methodological principles, which are made explicit, developed, revised and, on occasions, rejected or refuted, are hard to account for within the framework of SE. It is especially difficult, within the framework of SE, to account for principles which (a) have a quasi *a priori* role in specifying requirements theories must satisfy in order to be accepted, but which at the same time (b) make substantial *physical* assertions about the nature of the universe. AOE, on the other hand, predicts the existence of such principles, with just the features that have been indicated.

(8) *Verisimilitude*. The so-called problem of verisimilitude arises because physics proceeds from one false theory to another, thus rendering obscure what it can mean to say that science makes progress. Popper (1963, ch. 10 and Addenda) tried to solve this problem within the framework of SE but, as Miller (1974) and Tichy (1974) have shown, this attempted solution does not work. Not only does SE fail to justify the claim that theoretical physics makes progress; it fails even to say what progress in theoretical physics *means*.

AOE solves the problem without difficulty.

First, the fact that physics does proceed from one false theory to another, far from undermining physicalism, and hence AOE as well, is just the way theoretical physics must proceed, granted physicalism. For, granted physicalism, any theory, T^* , which captures precisely how phenomena evolve in some restricted domain, must be generalizable to cover *all* phenomena. If T^* cannot be so generalized then, granted physicalism, it cannot be precisely true. In so far as physics proceeds by developing theories which apply to restricted, but successively increasing, domains of phenomena, it is bound (granted physicalism) to proceed by proposing one false theory after another.

Second, AOE solves the problem of what it can *mean* to say that theories, T_0, \dots, T_N , get successively closer and closer to the true theory-of-everything, T , as follows. For this we require that T_N can be "approximately derived" from T (but not vice versa), T_{N-1} can be "approximately derived" from T_N (but not vice versa), and so on down to T_0 being "approximately derivable" from T_1 (but not vice versa).

The key notion of "approximate derivation" can be indicated by considering a particular example, the "approximate derivation" of Kepler's law that planets move in ellipses around the sun (K) from Newtonian theory (NT).

The "derivation" is done in three steps. *First*, NT is restricted to N body systems interacting by gravitation alone within some definite volume, no two bodies being closer than some given distance r . *Second*, keeping the mass of one object constant, we consider the paths followed by the other bodies as their masses tend to zero. According to NT , in the limit, these paths are precisely those specified by K for planets. In this way we recover the *form* of K from NT . *Third*, we reinterpret this "derived" version of K so that it is now taken to apply to systems like that of our solar system. (It is of course this *third* step of reinterpretation that introduces error: mutual gravitational attraction between planets, and between planets and the sun, ensure that the paths of planets, with masses greater than zero, must diverge, however slightly, from precise Keplerian orbits.)

Quite generally, we can say that T_{r-1} is "approximately derivable" from T_r if and only if a theory empirically equivalent to

T_{r-1} can be extracted from T_r by taking finitely many steps of the above type, involving (a) restricting the range of application of a theory, (b) allowing some combination of variables of a theory to tend to zero, and (c) reinterpreting a theory so that it applies to a wider range of phenomena.

This solution to the problem of what progress in theoretical physics *means* requires AOE to be presupposed; it does not work if SE is presupposed. This is because the solution requires one to assume (a) that the universe is such that a yet-to-be-discovered, true theory of everything, T , exists, and (b) current theoretical knowledge can be approximately derived from T . Both assumptions, (a) and (b), are justified granted AOE; neither assumption is justifiable granted SE.¹³⁵

(9) *Discovery of New Fundamental Theories.* Given SE, the discovery of new fundamental physical theories that turn out, subsequently, to meet with great empirical success, is inexplicable. (One thinks here of Newton's discovery of his mechanical theory and theory of gravitation, Maxwell's discovery of classical electromagnetism, Einstein's discovery of the special and general theories of relativity, Bohr's discovery of "old" quantum theory, Heisenberg's and Schrödinger's discovery of "new" quantum theory, Dirac's discovery of the relativistic quantum theory of the electron and, in more recent times, the discovery of quantum electrodynamics, the electroweak theory, quantum chromodynamics, the standard model and string theory.) Granted that a new theory is required to explain a range of phenomena, there are, on the face of it, infinitely many possibilities. In the absence of rational guidance towards good conjectures, it would seem to be infinitely improbable that anyone should, in a finite time, be able to come up with a theory that successfully predicts

¹³⁵. It may be objected that if T is assumed to be the true *unified* theory of everything, no meaning can be given to the idea that theoretical physics is making progress, by means of a succession of false theories, to a more or less *disunified* theory of everything. But T does not need to be assumed to be unified; all that is required is that T is such that the notion of "partial derivation" from T makes sense. For further discussion of the inability of SE to solve the problem of verisimilitude, and the ability to AOE to solve the problem, see Maxwell (1998), 70-72, 211-217 and 226-227.

new phenomena. Not only does SE fail to provide such rational guidance; the only guidance that it is able to offer is of exactly the wrong kind, and would, if heeded, lead one persistently astray. Granted SE, the only way to proceed, so it would seem, is to generalize and adapt existing theories to form new theories applicable to the new range of phenomena. But new fundamental theories often do not emerge in this way. They almost invariably contradict earlier theories, and often have a conceptual structure that differs fundamentally from predecessor theories, most notable in the transition from Newton's to Einstein's theories of gravitation, and in the transition from classical to quantum theory. It is not surprising that those who defend versions of SE tend to deny that the discovery of new theories is any kind of rational process.¹³⁶ But if it is not rational, it becomes miraculous that good new theories are ever discovered. Scientific progress becomes, in other words, inexplicable.

AOE, by contrast, provides physics with a rational, if fallible and non-mechanical, method for the discovery of new fundamental physical theories. This method involves modifying the current best level 3 blueprint so that:

- (a) the new blueprint exemplifies physicalism better than its predecessor;
- (b) the new blueprint promises, when made sufficiently precise to become a testable theory, to unify clashes between predecessor theories;
- (c) the new theory promises to exemplify the new blueprint better than the predecessor theories exemplify the predecessor blueprint.¹³⁷

¹³⁶. Reichenbach (1938, pp. 381-383), Popper (1959, p. 31), Kuhn (1970a) and Lakatos (1970) all, in effect, deny that discovery of new fundamental physical theories, new "paradigms" or "hard cores", is rational.

¹³⁷. Kuhn (1970a) gives a brilliant description of the way, during a scientific revolution, competing arguments for the rival paradigms are circular, each presupposing what is being argued for. But this can be seen to be a direct consequence of trying to do science in accordance with SE. This banishes explicit discussion of theses at levels 3 and 4; no wonder conservatives assess the new theory in terms of the old blueprint, and are unimpressed by the revolutionary assessment of the new theory in terms of the new blueprint. (The Gestalt switch, described by Kuhn, can be pin-pointed as the act of abandoning

(a), (b) and (c) provide means for assessing how good an idea for a new theory is which do not involve empirical testing (which is brought in once the new theory has been formulated). The level 4 thesis of physicalism provides continuity between the state of knowledge before the discovery of the new theory, and the state of knowledge after this discovery. Modifying the current blueprint ensures that the new theory will be incompatible with its predecessors; it will postulate new kinds of entities, forces, space-time structure, and will exhibit new symmetries. In other words, because of the hierarchical structure of AOE, there is (across revolutions) both continuity (at level 4) and discontinuity (at levels 2 and 3), something that is not possible given SE. AOE provides physics with specific non-empirical tasks to perform, specific non-empirical problems to be solved, and non-empirical methods for the assessment of ideas for new theories, all of which adds up to a rational, if fallible, method of discovery. It all stems from recognizing that physicalism is a part of current scientific knowledge. The discovery of new fundamental physical theories ceases to be inexplicable. None of this is possible granted SE.¹³⁸

The fact that AOE is able to provide a rational method of discovery, while SE is not, is due to the greater rigour of AOE (a point mentioned in (1) above). AOE has greater rigour than SE because AOE acknowledges, while SE denies, metaphysical assumptions implicit in persistent scientific preference for simple, explanatory theories. It is precisely the explicit acknowledgement of these metaphysical assumptions which makes the rational method of discovery of AOE possible.

(10) *Scientific practice*. Within the framework of SE, no one has succeeded in formulating a set of methodological rules which has won general acceptance as doing justice to scientific practice.

the old blueprint and adopting the new one.) Accept AOE, however, and it becomes possible for both parties, not only to acknowledge both old and new blueprints (distinct from corresponding theories), but also to discuss the adequacy of the new blueprint from the common standpoint of how adequately it exemplifies physicalism, which ought to be accepted by both parties. The irrationality of revolutions, depicted by Kuhn, disappears.

¹³⁸. For further discussion of the method of discovery provided by AOE see Maxwell (1974), Part II; (1993), Part III; and (1998), 159-163 and 219-223.

Instead, there are sustained, fundamental disagreements about the nature of scientific method. Inductivists stress that laws and theories must be arrived at cautiously by generalizing from observed phenomena. Popper denies that this ever happens, and argues that scientists must boldly speculate and seek to refute their speculations experimentally. Kuhn stresses that most of the time scientists protect the established theory or paradigm from refutation, normal science being devoted to showing that the established theory is able successfully to predict phenomena. Lakatos depicts science as consisting of competing research programmes, competing fragments of Kuhn's normal science. And Feyerabend concludes that none of the available views does justice to all good science, and that therefore the best policy is to hold that "anything goes". As Sokal and Bricmont (1998, p. 56) have recently put it "there does not exist (at least at present) a complete codification of scientific rationality, and we seriously doubt that one could ever exist".

By contrast, AOE is able to do justice to the range and richness of scientific practice, and is able, at the same time, to make clear why all SE attempts to depict scientific method fail. The fundamental natural science, theoretical physics, adopts a hierarchy of cosmological assumptions, to which correspond non-empirical methodological rules. As knowledge and understanding improve, cosmological assumptions and associated methods, lower down in the hierarchy, improve as well: more or less specific aims-and-methods improve within a framework of fixed aims-and-methods. Scientific method evolves with evolving knowledge. In order to do justice to this vital feature of scientific rationality, AOE characterizes scientific method at ten different levels. At level 1 there are observational and experimental methods, instruments and experimental techniques, associated with advances in empirical knowledge at level 1. At level 1, most strikingly, new knowledge can be used to acquire yet more new knowledge (as when, to give just one example of something that pervades the whole of science, Galileo uses the telescope to acquire astronomical knowledge). Here, too, knowledge generates knowledge. At level 2, new theoretical knowledge can be used to generate and assess further new knowledge, as when theory is used to predict new phenomena,

and call experimental results into question which clash with accepted theory. At level 3, there are (non-empirical) methods used to develop and assess level 2 theory, that are associated with the best available level 3 blueprint. There are meta-methods associated with the level 4 thesis of physicalism which govern, in part, acceptance of level 3 blueprints and associated methods. There are meta-meta-methods (meta²-methods) associated with the level 5 thesis of comprehensibility which govern, in part, acceptance of level 4 physicalism and associated meta-methods. And there are meta³-methods, meta⁴-methods, up to meta⁷-methods, associated with cosmological theses at levels 6 to 10, none of which would need to be invoked explicitly in science unless current scientific knowledge turned out to be dramatically on the wrong lines. This meta-meta structure enables AOE to do justice to the radical evolution of scientific method, from the Presocratics and Aristotle, via Galileo, Newton, Darwin, Einstein down to scientists at work today. At some level, there are (or ought to be) common assumptions of comprehensibility or knowability, and common associated methods; at lower levels there are diverse more specific assumptions and associated methods. To discern the unity in the diversity, and the rationale for the diversity, it is essential to adopt the meta-meta, hierarchical viewpoint of AOE.

This is also able to do justice to the diversity of methods to be found in diverse sciences, without underlying unity and rationality being sacrificed. It is important to appreciate, first, that different branches of the natural sciences are not isolated from one another: they form an interconnected whole, from theoretical physics to molecular biology, neurology and the study of animal behaviour. Different branches of natural science, even different branches of a single science such as physics, chemistry or biology, have, at some level of specificity, different aims, and hence different methods. But at some level of generality all these branches of natural science have a common aim, and therefore common methods: to improve knowledge and understanding of the natural world. All put aim-oriented empiricism into practice, but because different scientific specialities have different specific aims, at the lower end of the hierarchy of methods different specialities have somewhat

different methods, even though some more general methods are common to all the sciences. Furthermore, all natural sciences apart from theoretical physics presuppose and use results from other scientific specialities, as when chemistry presupposes atomic theory and quantum theory, and biology presupposes chemistry. The results of one science become a part of the presuppositions of another. This further enhances unity throughout diversity, and helps explain the need for diversity of method.

But in order to exhibit the rationality of the diversity of method in natural science, apparent in the evolution of methods of a single science, and apparent as one moves, at a given time, from one scientific speciality to another, it is essential to adopt the meta-meta, hierarchical standpoint of AOE, which alone enables one to depict methodological unity (high up in the hierarchy) throughout methodological diversity (low down in the hierarchy). SE, lacking this hierarchical structure, cannot begin to do justice to this key feature of scientific method, unity throughout diversity; nor can it begin to do justice to the rational *need* for this feature of scientific method.

Attempts to depict scientific method within the framework of SE are further cramped and distorted by the need to do justice to the point that, in physics, theories are only accepted if they accord sufficiently well with physicalism *without this being explicitly acknowledged*. As we saw above, SE must acknowledge that theory choice in physics is influenced by considerations of simplicity or explanatory capacity, but is unable to say what simplicity is, or why persistent preference for simplicity is justified.

There is a further, crucial point. Any new conception of science which improves our understanding of science ought to enable us to improve scientific practice. It would be very odd if our ability to do science well were wholly divorced from our understanding of what we are doing. A test for a new theory of scientific method ought to be, then, that it improves scientific practice, and does not merely accurately depict current practice. AOE passes this test. In providing a framework for the articulation and scrutiny of level 3 metaphysical blueprints, as an integral part of science itself, this providing a rational means for the development of new non-

empirical methods, new symmetry principles, and new theories, AOE advocates, in effect, that current practice in theoretical physics be modified. It makes explicit what is at present only implicit. And more generally, in depicting scientific method in a hierarchical, meta-meta fashion, AOE has implications for method throughout the natural sciences, and not just for theoretical physics.

In case it should seem miraculous that science has made progress without AOE being generally understood and accepted, I should add that good science has always put something close to AOE into practice in an implicit, somewhat covert way, and it is this which has made progress possible. The attempt to do science in accordance with the edicts of SE has been sufficiently half-hearted, sufficiently hypocritical, to make progress possible. It is only the fully rigorous implementation of SE that would have brought science to a standstill.

(11) *Aim-oriented empiricism uniquely equipped to explain unprecedented success of modern science.* The hierarchical structure of cosmological theses of AOE is designed to maximize the capacity of science to develop and choose those theses that best promote the growth of empirical knowledge, at levels 1 and 2. Any modification to the structure of AOE, which holds out a better hope of promoting such empirical growth, ought to be made. This means AOE is uniquely equipped to account for the rapid progress made by modern science. AOE is specifically designed to promote progress in scientific knowledge and understanding, however the universe may turn out to be. According to AOE, the key to scientific rationality is the positive feedback, engineered by AOE, that takes place between improving knowledge, and improving aims-and-methods, improving knowledge-about-how-to-improve-knowledge. No version of SE can incorporate this vital feature of AOE for maximizing the capacity of science to make progress.

7 Implications

What are the implications of pushing through the revolution that I have argued for, here, from SE to AOE?

First, there is a dramatic change in the whole relationship between science on the one hand, and metaphysics and philosophy

on the other. Given SE, metaphysics and philosophy are excluded from science, in accordance with Popper's criterion of demarcation: metaphysical theories (such as that the universe is physically comprehensible), being experimentally untestable, are unscientific. But granted AOE it is clear that untestable metaphysical or philosophical ideas are absolutely basic to scientific knowledge. Metaphysical theses at levels 4 to 10 in the diagram are more firmly established than currently accepted theories of physics, such as general relativity or quantum theory. No longer can philosophy be a forbidden subject for undergraduate physicists: on the contrary, it must be an important part of the curriculum!

But before it becomes a standard part of science in this way, philosophy must itself undergo a revolution. According to AOE, the proper way to assess metaphysical theories about the nature of the universe is in terms of their fruitfulness for science. This is not the way philosophers tend to assess such theories at present.

A second implication of adopting the new conception of science is that fundamental problems in the philosophy of science, unsolved for centuries, become readily resolved, as we have seen in the previous section.

A third implication of adopting the new conception of science is that science acquires a rational, if fallible and non-mechanical method for the discovery of fundamental new theories. In the previous section I concentrated on the most problematic case of the discovery of new fundamental theories within physics. But AOE, in stressing the need to make problematic *aims* (and problematic *assumptions* implicit in aims) explicit, has fruitful implications for discovery throughout science.

A fourth implication of adopting the new conception of science is that there is a dramatic change in the whole conception of scientific method. Within the framework of AOE, philosophy of science, the study of aims and methods, becomes a vital part of science itself, being shaped by, *but also helping to shape*, the way science evolves. This new, hierarchical, aim-oriented conception of scientific method has far reaching implications for rationality in general. For it is not just science that has problematic aims; our aims in life, whether individual, institutional or social, are

problematic. Above all, the aim of creating a better world is inherently and profoundly problematic. In these diverse fields, too, we need to put a generalized version of the progress-achieving methods of science into practice, designed to help us improve aims and methods as we proceed, as we live.¹³⁹

But finally, perhaps the most dramatic consequence of adopting the new conception of science is the following. Granted SE, the thesis that the universe is physically comprehensible is definitely not a part of current scientific knowledge. But granted AOE, this thesis is a central component of current theoretical knowledge in science, more firmly established, as I have said, than any accepted physical theory. This is implicitly, but not explicitly, recognized by physicists today when they concede that general relativity and the standard model, which do not form a unified theory, cannot therefore be correct. In holding that unity is a necessary condition for fundamental physical theory to be correct, physicists all but acknowledge that physicalism is a part of current knowledge.

They are prevented from acknowledging this explicitly by token allegiance to SE. The time has come to push through a revolution in our whole understanding of science. We need to reject standard empiricism in all its forms, and adopt aim-oriented empiricism in its stead as the new orthodoxy.

The argument of this paper reveals that there is a deep seated hypocrisy in humanity's understanding and use of science. In order to make sense of science we need to appreciate that science has discovered that the universe is physically comprehensible, in some way or other. This is however a discovery that is too disturbing for humanity readily to acknowledge. Scientists don't like to acknowledge the discovery, because it involves abandoning the idea that science differs radically from other enterprises – philosophy or religion – in that, in science, there are no permanent

¹³⁹. See Maxwell (1984 or 2007), (1992), (2000a), (2002a). The argument, as developed in these works, has far more radical implications for the nature of social inquiry than for the nature of natural science. Social inquiry becomes social *methodology* or social *philosophy*, rather than, primarily, social *science*. The philosophy of science, and the sociology of science, at present at odds with one another, become one and the same discipline: see Maxwell (1984), ch. 5, especially pp. 107-117. See also Maxwell (2002a).

assumptions upheld independently of empirical considerations, everything being decided impartially on the basis of evidence. Non-scientists (and many scientists) don't like to acknowledge the discovery because it involves accepting that we – all that we are, experience, think, feel, do – are embedded in, and are a part of, a physical universe governed by a fixed pattern of physical law. We use the products of science, but shield our eyes from the disturbing sight that science appears to reveal about our nature, our life, all that we hold to be of most value.¹⁴⁰

¹⁴⁰. For a discussion of the problem of how what is of value about our human world can be accommodated within a physically comprehensible universe see Maxwell (2001).

Chapter Six

Einstein, Aim-Oriented Empiricism¹⁴¹ and the Discovery of Special and General Relativity¹⁴²

(First published in the *British Journal for the Philosophy of Science* vol. 44, 1993, pp. 275-305)

According to Popper, Einstein is a falsificationist. Thus Popper declares: 'Einstein *consciously seeks error elimination*. He tries to kill his theories: he is *consciously critical* of his theories' (Popper (1972), p. 25). And elsewhere Popper declares: 'what I have done is mainly to make explicit certain points which are implicit in the work of Einstein' (Whitrow, 1973, p. 23). Paul Feyerabend, on the other hand, holds Einstein to be a methodological 'opportunist or cynic' or, in other words, a methodological anarchist (Feyerabend 1978, p. 213, note; see also p. 18, and pp. 56-7 and note). For Arthur Fine, Einstein adopts a view close to the natural ontological attitude (NOA). Fine writes: 'In its antimetaphysical aspect, NOA is at one with Einstein's motivational realism' (Fine, 1986, p. 9). As far as I know, van Fraassen has not yet claimed that Einstein is a constructive empiricist but, amazingly, the claim has been made on his behalf by Fine, who writes:

Indeed it would not be too far off if we summarized Einstein's views this way: 'Science aims to give us

¹⁴¹ See chapters 4, 5 and 10 for expositions of aim-oriented empiricism.

¹⁴² I am grateful to Harvey Brown for critical comments concerning the first draft of this section of the chapter.

theories which are empirically adequate; and acceptance of a theory involves as belief only that it is empirically adequate' [a straight quote from van Fraassen's *Scientific Image*] . . . My argument, then, is that if we understand Einstein in the way that he asks us to, his own realist-sounding language maps out a position closer to constructive empiricism than to either 'metaphysical realism' or 'scientific realism' (Fine 1986, p. 108).

The temptation to see one's own view in Einstein's thought is, it seems, all but irresistible. Do not I also give way to this temptation in attributing the views I have defended in earlier chapters of this book to Einstein?

I must confess that I did not arrive at these views as a result of reading Einstein. I developed aim-oriented empiricism during the course of criticizing Popper, and as the key to the solution to the problem of induction (see chapters 2 to 4, and my, 1979). At first I was convinced that standard empiricism had such a dogmatic stranglehold on science that it would be quite impossible for any scientist to uphold aim-oriented empiricism.¹⁴³ But it then began to dawn on me that Einstein, in developing special and general relativity, had made essential use of aim-oriented empiricism—his success owing much to his exploitation of the view in scientific practice.¹⁴⁴ I then discovered that Einstein had actually advocated

¹⁴³ For the distinction between aim-oriented empiricism and standard empiricism see chapter 4 or 5.

¹⁴⁴ It may be asked how it is possible for Einstein to be the first to exploit aim-oriented empiricism explicitly in scientific practice if what I have argued in part one of this paper is correct, and aim-oriented empiricism is inherent in all of science. The answer is straightforward. Actual scientific practice is massively influenced by the long-standing conviction of the scientific community that science ought to proceed in accordance with standard empiricism. The result is that scientific practice is a mixture of aim-oriented empiricism and standard empiricism. Aim-oriented empiricism is implemented in a surreptitious, hypocritical fashion, overlaid by the conviction that science ought to proceed in accordance with standard empiricism. As a result, physicalism and more specific metaphysical blueprints are not acknowledged within the intellectual domain of scientific knowledge, and this sabotages the possibility of

key tenets of aim-oriented empiricism in an increasingly explicit way as the years went by—but had been ignored and misunderstood because of the powerfully prevailing influence of standard empiricism.¹⁴⁵ Here are my reasons for holding this view.

Einstein invented aim-oriented empiricism in scientific practice in order to overcome a severe scientific crisis. The crisis was the demise of classical physics as a result of Planck's 1900 quantum theory of blackbody radiation. Initially, it was only Einstein who understood just how grave, how wholesale, the crisis was. In his 'Autobiographical Notes' he puts the matter like this.

it (became) clear to me as long ago as shortly after 1900, i.e. shortly after Planck's trailblazing work, that neither mechanics nor electrodynamics could (except in limiting cases) claim exact validity. By and by I despaired of the possibility of discovering the true laws by means of constructive efforts based on known facts. The longer and the more despairingly I tried, the more I came to the conviction that only the discovery of a universal formal principle could lead us to assured results. The example I saw before me was thermodynamics. The general principle was there given in the theorem: the laws of nature are such that it is impossible to construct a *perpetuum mobile* (of the first and second kind). How, then, could such a universal principle be found? (Einstein, 1969, pp. 51-3.)

putting the rational method of discovery of aim-oriented empiricism into sustained scientific practice. Explicit scientific exploitation of aim-oriented empiricism is frustrated if not prohibited. (See Maxwell [1976b and 1984] for further discussion of this point.) Einstein's great lucidity about fundamental matters led him to put aim-oriented empiricism into scientific practice unconstrained by hypocritical allegiance to standard empiricism.

¹⁴⁵ Gerald Holton comes the closest to interpreting Einstein in the way that I do. One difference, of course, is that Holton espouses his 'themata' conception of science and not aim-oriented empiricism: see Holton [1973].

This, I claim, is the beginning of the explicit employment of aim-oriented empiricism in scientific practice. It is to this that Einstein owed his extraordinary success in discovering special and general relativity. Soon after 1900, Einstein found himself bereft of guidelines as to how to proceed because Planck's 'trailblazing' result cast into doubt the whole of classical physics. Ordinarily a theoretical physicist can proceed by applying, extending, modifying or reinterpreting existing established physical theory. This is how classical physics had developed so far, after Newton. Einstein, however, found himself in what seemed an unprecedented situation. Existing physical theory—especially Newtonian mechanics and Maxwell-Lorentzian electrodynamics—must be fundamentally wrong, given Planck's result. A fundamentally new kind of theory was needed to stand in their stead. But, in order to discover this new theory, it would be useless to try to extend or modify existing physical theories, in the ordinary manner, since it was just these theories which were fundamentally wrong. In order to proceed Einstein was obliged to invent a new method of discovery for theoretical physics—a rational method capable of leading to the discovery of fundamentally new kinds of theories.

Within the framework of standard empiricism there can be no such rational method of discovery. If the *only* way in which theories can be rationally assessed in physics is by means of empirical success and failure, there can be no rational method for the invention of good, radically new physical theories which are incompatible with existing theories.

Popper (1959, 1963), Kuhn (1970a] and Lakatos (1970], all of whom defend versions of standard empiricism, not surprisingly all deny the possibility of there being a rational method of discovery of fundamentally new theories or paradigms—theories whose invention and acceptance constitute a 'scientific revolution'. Kuhn even denies that there can be rational *assessment* of a revolutionary new theory (with respect to its predecessor). The problem of how to proceed when confronted by wholesale scientific crisis, the breakdown of all existing theoretical knowledge, which Popper, Kuhn and Lakatos failed to solve in principle in the 1930s, 1960s and 1970s, Einstein had already solved in successful scientific

practice by the year 1905. He solved it by inventing special relativity.

What, then, is Einstein's new rational method of discovery, which led to the discovery of special and general relativity? It can be put, quite simply, like this. Choose two of the most fundamental physical theories, T_1 and T_2 say, which are a part of 'scientific knowledge' but which *contradict* each other. Discard everything about T_1 and T_2 that does not seem relevant to the contradiction until two mutually contradictory principles, P_1 and P_2 , are arrived at, P_1 from T_1 and P_2 from T_2 , thus arriving, it is hoped, at the *nub* of the contradiction between T_1 and T_2 . Modify P_1 or P_2 (or both) or relevant background assumptions so as to resolve the contradiction into a new unified principle, P_3 (a synthesis of a transformed P_1 and P_2). Take P_3 as the basis for a new theory T_3 , which unifies T_1 and T_2 .

In order for this method of discovery to be a rational one to adopt, one crucial assumption must be made: the universe has some kind of discoverable unified structure, of which our present fundamental physical theories give us limited, approximate (and incompatible) glimpses. Given the truth of this assumption, we have rational grounds for holding that the method can lead to success. If the assumption is false, we have no such grounds. As we shall see, Einstein seems only to have fully understood this point after the discovery of general relativity.

As far as the discovery of special relativity is concerned, Einstein used the above method in the following way. The two fundamental physical theories that he takes as his starting point (T_1 and T_2) are Newtonian mechanics (NM) and Maxwellian electrodynamics (ME). These two theories are incompatible, fundamentally because, given their most natural interpretation, NM is about forces-at-a-distance between point-particles with mass, whereas ME is about one entity, the continuous electromagnetic field. More specifically, however, there is the following contradiction. NM asserts that forces affect *accelerations*, not *velocities*. Dynamic laws (laws concerning forces and their affects), formulated within the framework of NM, do not pick out any special *velocity* any more than they pick out some special *place* or *time*. ME does, however, pick out a special velocity: the velocity of light, the

velocity that, according to ME, vibrations in the field strengths of the electromagnetic field travel through space.

Both points are absolutely fundamental to the two theories. It is fundamental to the whole structure of NM that forces affect accelerations, not velocities (there thus being no role for absolute velocity within the theory). And it is fundamental to ME that influences should spread through the field at some fixed, *finite* velocity: for it is this which creates the need for a field theory in the first place. (Because gravitational influences, in Newton's theory of gravitation, spread at infinite velocity, instantaneous physical states can be specified in terms of point-particles. When influences travel at some *finite* velocity, as in ME, this can no longer be done, as momentum and energy associated with variations in the force travelling at finite velocity through space will not be specified.)

One way in which the clash between NM and ME may be resolved is to interpret ME as a theory which presupposes the existence of the aether, states of the electromagnetic field being states of the aether. In this case, it is reasonable to hold that light has a constant velocity with respect to the aether, and the clash with NM disappears (the constancy of the velocity of light being as unproblematic as the constancy of the velocity of sound with respect to air). In his 1905 paper expounding special relativity, Einstein gave two reasons for rejecting this approach. First, it introduces an implausible asymmetry in the explanation of electromagnetic induction, implausible because of the symmetry in the phenomena to be explained. The theoretical explanation for the current in a conductor moving near a magnet at rest is strikingly different from the explanation of the current if the conductor is at rest and the magnet moves, even though all that matters is the relative motion as far as the effect is concerned. Second, it runs into empirical difficulties in that all attempts to detect the motion of the earth relative to the 'light medium'—the aether—have failed. Einstein concluded that 'the phenomena of electrodynamics as well as of mechanics possess no properties corresponding to the idea of absolute rest' (Einstein, 1905; translated in Einstein *et al.*, 1952, p. 37).

Einstein was of course well aware that the null result of the Michelson-Morley experiment does not decisively demolish the aether; he knew of Lorentz's efforts to employ the FitzGerald contraction hypothesis to develop a version of electrodynamics which both presupposes the aether and is compatible with observation. In a paper published in 1907, however, Einstein remarked of the FitzGerald-Lorentz approach (surely with some justice) that it is 'ad hoc' and 'artificial' (Holton, 1973, p. 334)—although, as Grünbaum and Zahar remind us, this approach is not as grossly *ad hoc* as some have supposed (Grünbaum, 1963, pp. 386-94; Zahar, 1973).

We know that during the decade before 1905, Einstein took the aether hypothesis sufficiently seriously to wonder how motion through the aether might be detected (Paris, 1982, pp. 130-2). Nevertheless, it seems that, early on, Einstein was drawn to what may be called the 'Faraday interpretation' of electromagnetism, according to which, instead of seeking to interpret electromagnetism in terms of some more fundamental kind of aetherial matter, one should, on the contrary, seek to understand matter in terms of electromagnetism, which is to be regarded as fundamental (the whole idea of the aether being a mistake). This is implicit in the 'paradox' that Einstein discovered when sixteen, and which he later saw as the germ from which special relativity grew. In his 'Autobiographical Notes', Einstein describes the paradox thus:

If I pursue a beam of light with the velocity c ... I should observe such a beam of light as a spatially oscillatory electromagnetic field at rest. However, there seems to be no such thing, whether on the basis of experience or according to Maxwell's equations. From the very beginning it appeared to me intuitively clear that, judged from the standpoint of such an observer, everything would have to happen according to the same laws as for an observer who, relative to the earth, was at rest. For how, otherwise, should the first observer know, i.e. be able to determine, that he is in a state of fast uniform motion? (Einstein, 1969, p. 53.)

This only makes intuitive sense as a paradox in so far as electromagnetism is being conceived of in the absence of the aether.

As I have argued elsewhere (Maxwell, 1979, pp. 647-8; 1988, p. 42; 1998, p. 80-82), there are strong aim-oriented empiricist, quasi *a priori* grounds for favouring the Faraday interpretation of electrodynamics over the aether interpretation, the view that the aether is required to make electrodynamics intelligible being a sort of metaphysical blunder. And there is an additional consideration. According to aim-oriented empiricism, the acceptability of the aether hypothesis is to be judged in terms of its heuristic and methodological power. But as ME was developed, up to 1905, especially in the hands of Lorentz, the role of the aether seemed to become increasingly tenuous. This, according to aim-oriented empiricism, counts against the aether approach. We may thus detect, in Einstein's adoption of the Faraday interpretation of electrodynamics, and his rejection of the aether interpretation, an instinctive allegiance to aim-oriented empiricism.

There is, however, another approach to resolving the clash between NM and ME. It is possible that the velocity of light is constant with respect to the source. Einstein tried this approach; he tells us that he abandoned it because of the complications to which it led (Shankland, 1963). Evidence against this hypothesis only began to come in later, in 1913, with observations of double stars.

Granted, then, that the above two approaches to resolving the clash between NM and ME are to be rejected, we are left with the following situation: ME appears to be committed to the existence of a fundamental, *absolute* velocity—the velocity of light—just that which NM rules out. We have here, then, two good candidates for P_1 and P_2 , extracted from T_1 (NM) and T_2 (ME) in order to highlight the clash between the two theories, namely:

P_1 : The laws of nature have the same form with respect to all inertial (non-accelerating) reference frames.

P_2 : It is a law of nature that light travels with constant velocity c (in a vacuum).

P_1 and P_2 together form, it would seem, a horrible contradiction. In order for P_1 and P_2 to be compatible it would be necessary for a beam of light to have the same velocity c with respect to *all* inertial

reference frames, even though these are moving with all possible velocities with respect to each other.

Astonishingly, Einstein discovered how to make this apparently blatant absurdity entirely consistent. What we need to do is to modify our ideas about time and space, so that light *does* have the same velocity c in all reference frames. The basic postulates of special relativity are just P_1 and P_2 : the many consequences of the theory arise from demanding that P_1 and P_2 be consistent.

More precisely, Einstein took P_1 as one of his basic postulates, but modified P_2 to become:

P_2^* : It is a law of nature that the velocity of light is a constant c in some 'resting' reference frame, and is independent of the velocity of the source.

P_2 is then derived from P_1 and P_2^* . It is entirely understandable that Einstein took P_2^* as his axiom rather than P_2 interpreted to mean: It is law of nature that light has constant velocity c in all inertial reference frames. To adopt this latter postulate is to assume as comprehensible that which only becomes comprehensible with the development of the theory. P_2^* is not initially incomprehensible in this way; on the contrary, P_2^* is a basic tenet of the Lorentzian approach, of the aether approach widely held at the time.

How, then, is the contradiction between P_1 and P_2 to be resolved? Ordinarily we assume that the rate of clocks, and the length of rods, are unaffected by uniform motion, temporal and spatial distances being frame-independent and absolute. Suppose we have two reference frames, R_1 and R_2 , with parallel axes, and with origins that coincide at $t_1 = t_2 = 0$, the origin of R_2 travelling along the x axis of R_1 with velocity v in the +ve direction, the coordinates of an event P being (x_1, y_1, z_1, t_1) and (x_2, y_2, z_2, t_2) in R_1 and R_2 respectively. In effect, we ordinarily assume that the coordinates are related by the 'Galilean' transformations:

$$x_2 = x_1 - vt_1; \quad y_2 = y_1; \quad z_2 = z_1; \quad t_2 = t_1.$$

We assume, that is, that length and time are unaffected by motion, and that if a pulse of light which has velocity c along the x axis in the +ve direction in R_1 , then its velocity in R_2 is $c - v$.

What Einstein realized was that if rates of clocks and lengths of rods *are* affected by relative motion, so that $x_2 = x_1 - vt_1$

and $t_2 = t_1$ are both *false*, then it is entirely possible that any given pulse of light has the same velocity c in *both* R_1 and R_2 — indeed, the same velocity c in *all* inertial reference frames.

It turns out that the thesis that light does have the same velocity in all inertial reference frames — which is implied by P_1 plus P_2 — suffices to fix uniquely just how the coordinates of R_1 and R_2 are related. All that we need, in addition, is that the relationship is *symmetric* (which may be said to be inherent in P_1 in any case), *linear* and *isotropic*. With these assumptions it is not hard to show that the coordinates of R_1 and R_2 are related by the following equations, the 'Lorentz' transformations:

$$\frac{x_2 = x_1 - vt_1 \quad y_2 = y_1 \quad z_2 = z_1}{\sqrt{1 - v^2/c^2}} \quad \frac{t_2 = t_1 - vx_1/c^2}{\sqrt{1 - v^2/c^2}}$$

According to these equations, all but uniquely determined by P_1 plus P_2 , relative motion contracts rods and makes clocks go slow, but in such a way that the velocity of light is c in *all* inertial frames. The miracle of reconciling P_1 and P_2 has been achieved.

Special relativity has a number of startling implications. One is that *mass*, along with the speed of clocks and length of objects, is affected by uniform motion, so that $m_2 = m_1/(1 - v^2/c^2)^{1/2}$, where m_1 is mass of object in rest frame R_1 , with respect to which the object is at rest, and m_2 is mass of object in R_2 . Another—the most famous of all—is that *mass* is a form of *energy*, in accordance with $E = mc^2$.

From the standpoint of aim-oriented empiricism, special relativity is doubly significant. First, the way in which Einstein *discovered* special relativity exemplifies the method of discovery of aim-oriented empiricism, to the extent that Einstein used the method I have indicated above, namely: create a new theory as the outcome of resolving a clash between two existing theories— thus creating greater conceptual and theoretical unification. Secondly, and quite strikingly special relativity itself exemplifies aim-oriented empiricism, and in an important sense cannot be adequately understood within the framework of standard

empiricism. For, as Einstein himself remarks, two pages on from the quotation given above from his 'Autobiographical Notes':

The universal principle of the special theory of relativity is contained in the postulate: The laws of physics are invariant with respect to the Lorentz-transformations . . .

This is a restricting principle for natural laws, comparable to the restricting principle of the non-existence of the *perpetuum mobile* which underlies thermodynamics (Einstein, 1969, p. 57).

Special relativity is thus a law of laws, a meta law, a guiding principle, a heuristic and methodological rule to be employed in discovering and assessing physical theories—above all, for Einstein of 1905, to be employed as a heuristic tool for the discovery of the new theory to unify classical mechanics and electrodynamics. (When viewed from this perspective, what Einstein did in creating special relativity was to take a basic restricting principle of Newtonian mechanics, namely Galilean invariance—the pre-relativistic way to interpret P_1 —and modify this to make it compatible with P_2 , thus forming a new restrictive principle, P_3 , i.e. Lorentz invariance.) As a heuristic and methodological principle, special relativity has amply fulfilled Einstein's hopes for it. It played a vital role in the discovery of de Broglie's wave theory of matter, the so-called Klein-Gordon equation (first discovered by Schrödinger), the Dirac equation of the electron, quantum electrodynamics, quantum electro weak theory, and quantum chromodynamics. In a modified form, it played a crucial role in the discovery of general relativity; and it continues to be relevant to superstring theory. Here, then, is a heuristic and methodological principle of enormous fruitfulness for all of theoretical physics, which can be formulated as the demand that acceptable theories must be Lorentz invariant. This demand—equivalent to the demand that space-time be Minkowskian (in the formulation of theories)—is not *merely* a methodological principle for, as we have seen, it has substantial physics in it. Special relativity is capable of being falsified and, from the standpoint of general relativity, it is false. All this is very hard to make sense of, or do justice to, within the confines of any version of standard empiricism, precisely because standard empiricism *rejects* the idea

that methodological principles have physics, or metaphysics, built into them— there being, within standard empiricism, no (level 2) *metamethodological* framework within which rival (level 1) *methodological* principles can be rationally assessed. From the standpoint of aim-oriented empiricism, there is no difference in principle between an *ordinary* methodological rule such as *position invariance* (acceptable laws and theories must be invariant with respect to change of position in space), and full *Lorentz invariance*. To both there correspond substantial *physical* or *metaphysical* principles, namely: 'space selects out no special position' or 'space-time selects out no special inertial reference frame'. Both may be false, and both therefore require critical scrutiny as science develops, in accordance with aim-oriented empiricism, and not dogmatic acceptance or rejection, as required by standard empiricism, with its *fixed* set of methodological principles (which no one yet has been able to formulate!). Standard empiricism differentiates sharply between the status of *position* and *Lorentz invariance*—only the former qualifying as a methodological rule of physics, the latter belonging exclusively to the *content* of physics, as a physical theory. But this does violence to Einstein's achievement; it does violence to the new way of doing physics inspired by Einstein, which precisely *exploits* the fruitful interplay between new theories and new heuristic and methodological principles (along the lines stipulated by aim-oriented empiricism).

Aim-oriented empiricism is even more explicit in Einstein's discovery of general relativity. Einstein exploits the same method of discovery. As before, there are two fundamental conflicting theories, namely: Newton's theory of gravitation (T_1) and special relativity (T_2). These conflict because whereas Newton's theory implies that gravitational influences travel instantaneously, special relativity implies that such influences cannot travel faster than light. As before, Einstein searches for new *principles* which will guide him to a new unifying theory. His first step is to notice that there is a principle implicit in Newton's theory of gravitation (P_1) which, if generalized (P_1^*), makes it possible to generalize and *improve* the principle of relativity basic to special relativity (P_2). This latter principle seemed unsatisfactory to Einstein because of its restriction to some arbitrarily selected set of inertial reference

frames all in uniform motion with respect to each other. Much more satisfactory would be a *general* principle of relativity (P_2^*) which asserts that the laws of nature have the same form in *all* reference frames, however they may be moving or accelerating with respect to each other. But this general principle of relativity seems impossible to implement. It is one thing to say, given a train moving uniformly through a station, that there are two equivalent descriptions: (1) train moving, platform at rest; and (2) train at rest, platform moving (in opposite direction). It is quite another to say, given that the train crashes into the buffers at the end of the station, that there are two equivalent descriptions: (1) train de-accelerates, platform remains unaccelerated; and (2) train remains unaccelerated, platform de-accelerates. These are not equivalent descriptions: in the first, it is people in the train that suffer from violent de-acceleration, whereas in the second it is people on the platform that suffer. But consider now the following remarkable feature of Newton's law of gravitation (P_1): in a uniform gravitational field all objects accelerate equally, whatever their mass (essentially because inertial and gravitational mass are equal). Generalize this to form the *principle of equivalence* (P_1^*): no local phenomenon distinguishes between (a) uniform acceleration, and (b) being at rest in a uniform gravitational field. Whatever effect a gravitational field has on some phenomenon, it is the same as the effect that the equivalent acceleration would have in the absence of gravitation. This immediately has two consequences. First, it allows us to hold that *all* frames, however accelerating, are equivalent, as long as, in moving from one frame to another accelerating with respect to the first, we can invoke an additional, compensating gravitational field. Thus, in the case of the crashing train we have: (1) train de-accelerates, platform remains stationary; (2) train remains stationary, platform de-accelerates, and a gravitational field exists momentarily to compensate precisely for this de-acceleration. In *both* cases, it is the people in the train who suffer, according to the first description, because of de-acceleration, according to the second, because of the sudden gravitational field (and no compensating de-acceleration, as on the platform). The generalized principle of equivalence (P_1^*) makes it possible, in this way, to hold the

generalized principle of relativity (P2*). The second consequence of the generalized principle of equivalence (P1*) is that, if correct, it enables us to discover the effects that uniform gravitational fields have on phenomena; all we need to do is to consider the effects of uniform *acceleration* and put these equal to the effects of the corresponding gravitational field in the absence of acceleration. The principle of equivalence (P1*) thus has great potential heuristic power for the discovery of the new theory of gravitation, to replace Newtonian theory.

According to special relativity, acceleration affects geometry. Consider a flat, rapidly rotating disk. A rigid rod, of length L at the centre of the disk will, according to special relativity, only have length $L(1 - v^2/c^2)^{1/2}$ at the circumference, given that it is aligned with the motion of rotation which, at the circumference, has the value v . The geometry of the disk, as determined by the rod, will be non-Euclidean.¹⁴⁶ Uniform circular motion is accelerated motion. But if acceleration affects geometry, so, too, by the principle of equivalence, must gravitation. We have the possibility that gravitation *is* the (non-Euclidean) curvature of space-time—a possibility which, if true, would bring about a tremendous conceptual unification in the foundations of physics (namely the unification of gravitation and space-time geometry). Postulate therefore that gravitation is indeed the curvature of space-time. The presence of matter curves space-time; and matter moves along geodesics in this curved space-time. Curved space-time can always be reduced to flat Minkowskian space-time in any infinitesimal region by an appropriate choice of coordinate system, in accordance with the principle of equivalence given its final local formulation (P1**). What remains to be done is to formulate the precise way in which energy-momentum affects the Riemannian curvature of space-time.¹⁴⁷ The field equations of general relativity are the simplest possible solution to this problem. Indeed, granted

¹⁴⁶ For a more detailed discussion of the role played by the rotating disk in the genesis of general relativity, see Stachel [1980].

¹⁴⁷ For a discussion of this part of Einstein's creation of general relativity, see Pais ([1982], Chs. 11 and 12), or, even better, Norton [1984].

that the equations involve derivatives no higher than the second, the field equations are determined uniquely to be:

$$R_{ab} - 1/2g_{ab}R = 8\pi GT_{ab}$$

Here R_{ab} is the Ricci tensor of the metric g_{ab} (the Ricci tensor being derivable from the Riemannian curvature tensor by contraction), R is the Ricci scalar (formed from R_{ab} by contraction), T_{ab} is the energy-momentum tensor, and G is Newton's constant of gravitation.¹⁴⁸ We have arrived at T_3 , which reduces to special relativity (T_2) in the absence of gravitation, and which approaches Newtonian theory (T_1) in the limit as gravitational fields become weak and velocities become low in comparison with the velocity of light.

Does Einstein really put aim-oriented empiricism into practice in developing the special and general theories of relativity, in the way I have just sketched? Is there, here, a genuine method of discovery, given that Einstein failed for over thirty years to develop a satisfactory unified field theory? A few comments are in order.

The full aim-oriented empiricist method of discovery involves the tackling of at least four kinds of problems: (1) conflicts between experimental results and theory; (2) conflicts between well-established fundamental theories; (3) conflicts between such theories and the best available blueprint¹⁴⁹ for physics; and (4) conflicts inherent in the best blueprint itself (or between rival blueprints). It could be argued that Einstein only exploits a small part of this method of discovery, in that he is primarily concerned with type (2) problems (and type (1) problems where relevant). But this is, I think, wrong for a number of reasons.

¹⁴⁸ Good expositions of general relativity are to be found in Friedman [1983], Schutz [1988] and Misner *et al.* [1973].

¹⁴⁹ The best "blueprint" for physics is the best available idea as to how the universe is physically comprehensible, a vital element of theoretical scientific knowledge according to aim-oriented empiricism: see chapters 3, 4, 6 and 7.

First, the metaphysical thesis that the basic laws of nature have a *unified* structure is an implicit or explicit assumption in all of Einstein's deliberations.

Second, in developing special and general relativity it is precisely the pre-existing metaphysical blueprints of classical physics which Einstein is led to transform—basic assumptions about the nature of space, time, energy, mass, force. In developing new *principles*—such as the principle of Lorentz invariance or the principle of equivalence—Einstein is, at one and the same time, modifying pre-existing blueprint ideas (Newtonian space-time being transformed into Minkowskian space-time which is in turn transformed into the Riemannian space-time of general relativity).

Third, lurking behind the type (2) problems which concern Einstein (involving clashes between *theories*) there are type (4) *blueprint* problems. Consider the type (2) problem that led to special relativity—the clash between Newtonian mechanics and Maxwellian electrodynamics or, more specifically, the clash between Galilean invariance and the thesis that the constancy of the velocity of light is a law of nature. Around 1900, as we have seen, there was an obvious solution to this problem: interpret electrodynamics in terms of the aether, regard the constancy of the velocity of light as being relative to the aether, and expect Galilean invariance to break down for high velocities with respect to the aether. This amounts, of course, to adopting a blueprint for physics—the aether blueprint. In formulating the problem in the way in which he did, Einstein is in effect *rejecting* this aether blueprint; he is adopting Faraday's view that the field is fundamental, and does not require an underlying aether to make it comprehensible. As I have argued elsewhere (Maxwell, 1979, pp. 647-8, 1988, p. 42; 1998, p. 80-82), there are good reasons for preferring what may be termed the Faraday blueprint to the aether blueprint. The important point, however, is that in formulating his type (2) problem in the way in which he did (crucial for the development of special relativity), Einstein is in effect interpreting Newton's and Maxwell's theories to be two equally fundamental, rival theories, each with its rival *blueprint*, namely, the Newtonian (or Boscovichean) blueprint of point-particles surrounded by spherically symmetrical, rigid fields, and the Faraday field

blueprint with variations in the field being transmitted at some *finite* velocity. There is, in short, a type (4) blueprint problem inherent in the type (2) problem which led Einstein to special relativity. This type (4) problem may be formulated, not as a problem about how to reconcile, or choose between, two rival blueprints, but rather as the problem of how to resolve the clash that results from attempting to unify the two blueprints in such a way as to accommodate charged point-particles *and* a field.

Fourth, there are grounds for holding that Einstein's fundamental problem soon after 1900 was the type (4) *blueprint* problem I have just indicated—the problem of understanding how charged point-particles can interact with the field, or the problem of unifying point-particle and field. It is a striking fact that Einstein's three great papers of 1905 can all be interpreted as exploring aspects of this fundamental problem. We have just seen that this is true of the paper introducing special relativity. It is also true of the Brownian motion paper, concerned to establish the existence of atoms—the existence of the particle-like aspect of reality. And it is true above all of the paper which put forward the idea that light has a particle-like aspect in accordance with $E = nh\nu$ (where E is the energy and ν the frequency of the light, h is Planck's constant and n is some integer, the number of light quanta present), this 'heuristic' hypothesis of light quanta then being used to explain the photoelectric effect. Here the classical particle/field problem is intensified to an extraordinary extent in that the field itself is revealed to have a particle-like aspect.¹⁵⁰

As it happens, Einstein himself makes clear in his 'Autobiographical Notes' (Einstein, 1969) that he held the classical particle/field problem to be of fundamental importance.¹⁵¹ Having explained that theories are to be critically assessed from the two

¹⁵⁰ Lucid summaries of these papers are to be found in Lanczos [1974]. They are reproduced, translated into English, in Stachel [1998].

¹⁵¹ He did this too in the introduction to his 1905 paper introducing the revolutionary idea of light quanta. He there makes clear that there is a fundamental clash between the idea of the particle, associated with theories of matter, and the field idea of Maxwell's theory of electrodynamics: (see Stachel [1998], pp. 177-8).

distinct standpoints of empirical success and 'inner perfection' (unity or comprehensibility)—which in itself commits Einstein to aim-oriented empiricism (see below)—he goes on to assess critically Newtonian mechanics and Maxwellian electrodynamics from the standpoint of inner perfection. We have here, incidentally, an adjunct to, and refinement of, Einstein's method of discovery: *one* theory is here taken at a time, and is assessed from the standpoint of 'inner perfection'—from the standpoint, that is, of the capacity of the theory to provide a 'perfect' blueprint for all of physics as far as the *form* of the theory is concerned. (In indicating the 'inner perfection' *defects* of a theory one in effect indicates, at least in general terms, what would constitute a 'perfect' theory: one indicates, that is, a blueprint.) Einstein discusses five 'inner perfection' defects in Newtonian mechanics, namely: (1) arbitrariness in the determination of inertial reference frames from an infinity of alternatives, and inadequacy of introducing absolute space (with respect to which all bodies have absolute acceleration as a solution to this problem); (2) two distinct basic laws (and not *one*), namely: (a) the law of motion ($F = ma$), and (b) the expression for force or potential energy ($F = Gm_1m_2/d^2$): (3) arbitrariness of (b) given (a), there being endlessly many equally good possibilities for (b) given (a); (4) the possibility of the force law being determined by the structure of space (the form of the force law being suggestively simple when viewed in geometrical terms), and yet the *failure* to exploit this possibility; (5) the *ad hoc* character of the equality of inertial and gravitational mass; and (6) unnaturalness of energy being split into two forms, kinetic and potential (see Einstein, 1969, pp. 27-31). As far as electrodynamics is concerned, Einstein discerns one basic defect associated with interpreting the field equations as applying to matter and, in the case of the vacuum, to the aether. Einstein argues (perhaps not altogether accurately) that this defect was overcome by Lorentz in reinterpreting the field equations to hold essentially only for the vacuum, with matter, in the form of charged particles, being the source of the field. Einstein then remarks: 'If one views this phase of the development of theory critically, one is struck by the dualism which lies in the fact that the material point in Newton's sense and the field as continuum are used as elementary concepts

side by side' (Einstein, 1969, p. 37). Einstein explains why attempts to overcome this basic defect by eliminating the point-particle do not succeed; and he concludes: 'Accordingly, the revolution begun by the introduction of the field was by no means finished. Then it happened that, around the turn of the century ... a *second fundamental crisis* set in' (my italics)—namely the crisis engendered by the first step towards quantum theory, Planck's quantum explanation of his empirical radiation law. If this is the *second* fundamental crisis, then the first is particle/field dualism of classical physics. As it happens, the two crises are intimately interrelated, since Planck's law and quantum theory deal with the *interaction* of field and matter.

There are good grounds, then, for holding that Einstein was concerned with problems from type (2) to type (4), as defined above, and type (1) problems where relevant, a type (4) problem of special concern to Einstein being the problem of how to unify point-particles and field.

But did Einstein really invent an authentic method of discovery in view of his failure, during the last thirty years of his life, to discover the unified field theory he so ardently sought?

One reply can be made immediately: the method of discovery, indicated above, though rational, is also non-mechanical and fallible. The failure of the method to lead to a good fundamental new theory over a period of thirty years—even in the hands of Einstein—does not prove that the method is inauthentic.

But there is a much more important reply to be made. Einstein did not use his method of discovery in seeking to formulate his unified field theory. Or rather Einstein *misapplied* this method, in a quite elementary way.

After around 1930, the two fundamental theories that stand in most glaring contradiction with each other are general relativity and quantum theory. In order to implement Einstein's rational method of discovery from about 1930, the first step to take is to extract basic principles, P_1 and P_2 , from general relativity and quantum theory respectively, which *contradict* each other—this even perhaps being the *nub* of the contradiction between the two theories. The task then is to modify P_1 and P_2 (or something else)

to form P₃, a new principle which guides us to a new unified theory T₃, unifying general relativity and quantum theory.¹⁵²

Einstein did none of this. Instead he took as his two theories general relativity and classical electrodynamics and sought to unify these two theories, to form a theory which applied to all phenomena, including quantum phenomena. One may well have doubts as to whether these two theories really do fundamentally contradict each other—even though the theories are clearly *two* distinct theories and not *one* unified theory. They are at least both field theories; they both incorporate Lorentz invariance, at least locally; and they are both classical and deterministic. What is dramatically apparent is that the fundamental contradiction of theoretical physics after 1930 concerns, not the clash between classical general relativity and classical electrodynamics, but rather the clash between general relativity and quantum theory. (One can add that it is perverse to continue to take the unification of gravitation and electromagnetism as *the* unification to strive for, sufficient to create the comprehensive unified field theory, after the discovery of the strong and weak forces in addition to the forces of gravitation and electromagnetism.)

Why did Einstein so crudely and wilfully *misapply* his rational method of discovery? The answer is straightforward: because of his abhorrence of quantum theory given its orthodox interpretation (OQT). Einstein was absolutely correct to find OQT fundamentally defective from the crucial standpoint of 'inner perfection'. As I shall argue in a moment, Einstein's attitude towards OQT exemplifies yet again his (sound) commitment in scientific practice to aim-oriented empiricism and scientific realism. Where Einstein went *wrong* was to conclude that quantum theory was therefore entirely devoid of heuristic value—that it 'offers no useful point of departure for future development' (Einstein, 1969, p. 87).

¹⁵² Elsewhere, I have suggested that the way to implement this method of discovery is to take, as pi, the general idea of deterministic dynamic space-time geometry (from general relativity), and to take, as P2, the general idea of ontological probabilism (from quantum theory), the task then being to create unified probabilistic dynamic geometry, P3: (see Maxwell [1985], pp. 40-1).

What is striking about this is that it is actually a vital feature of Einstein's method of discovery that one deals with theories that are intrinsically *defective*. The defects are clues as to how the theory may be fruitfully modified. As we have seen above, Einstein indicates a number of fundamental defects inherent in Newtonian mechanics and Maxwellian electrodynamics. Einstein even knew, by 1901, as a result of Planck's work, that both theories are fundamentally incorrect. This did not stop him taking these theories as 'points of departure'. Indeed, it is the *defects* in the theories, as perceived by Einstein, which make his method of discovery so successful: for it is these defects which indicate how the theories are to be modified to overcome the contradictions between them. For Einstein to argue, after 1930, that the *defective* character of quantum theory ensures that the theory cannot form a proper point of departure does violence to the very heart of Einstein's own earlier method of discovery, used in the discovery of special and general relativity with such striking success.

Why did Einstein fail to recognize the fairly obvious point just made? In essence, because his abhorrence of OQT was so intense, so profound, that it was emotionally impossible for him to work seriously with the theory. He did not want to contribute to what he interpreted as a sickness which had entered physics, and which he regarded as symptomatic of the basic sickness of our times. In a sense, Einstein turned his back on quantum theory, and devoted himself to the task of unifying general relativity and classical electromagnetism as a kind of moral protest against the tenor of our times.

In order to substantiate this point I must now break off my discussion of Einstein's successes and failures in implementing aim-oriented empiricism so that I can consider in a little more detail the question of Einstein's attitude to OQT.

His mature attitude can be summarized like this. From the standpoint of empirical criteria, OQT must be judged to be an immense success. From the equally important standpoint of criteria having to do with 'inner perfection', with unification, OQT must be judged to be a disaster. This is because the theory cannot be interpreted to be about some hypothetical *reality*. It was not so

much the lack of *determinism* that came to worry Einstein as the lack of *realism*. In his 'Autobiographical Notes' he puts it like this.

Physics is an attempt conceptually to grasp reality as it is thought independently of its being observed. In this sense one speaks of 'physical reality'. In pre-quantum physics there was no doubt as to how this was to be understood. In Newton's theory reality was determined by a material point in space and time; in Maxwell's theory, by the field in space and time. (Einstein, 1969, pp. 82-3.)

Einstein goes on to point out that as far as OQT is concerned, there is no quantum equivalent to the classical material point or field. OQT makes probabilistic predictions about the results of performing measurements on an ensemble of similarly prepared systems, but cannot be interpreted as specifying the physical state of the individual system as it evolves in space and time independent of measurement. As Einstein puts it in volume 2 of the same book, in his 'Reply to Criticisms':

What does not satisfy me . . . [about OQT], from the standpoint of principle, is its attitude towards that which appears to me to be the programmatic aim of all physics: the complete description of any (individual) real situation (as it supposedly exists irrespective of any act of observation or substantiation) (Einstein, 1969, p. 667).

In a letter to Schrödinger in 1950, Einstein expresses himself even more emphatically.

You are the only contemporary physicist, besides Laue, who sees that one cannot get around the assumption of reality—if only one is honest. Most of them simply do not see what sort of risky game they are playing with reality—reality as something independent of what is experimentally established. They somehow believe that the quantum theory provides a description of reality, and even a *complete* description; this interpretation is, however, refuted, most elegantly by your system of radioactive atom + Geiger counter + amplifier + charge of gun powder + cat in a box, in which the ψ -function of the system contains the cat both alive and blown to bits. Is the state of the cat to be created only when a physicist

investigates the situation at some definite time? Nobody really doubts that the presence or absence of the cat is something independent of the act of observation. But then the description by means of the ψ -function is certainly incomplete, and there must be a more complete description. If one wants to consider the quantum theory as final (in principle), then one must believe that a more complete description would be useless because there would be no laws for it. If that were so then physics could only claim the interest of shopkeepers and engineers; the whole thing would be a wretched bungle (Przibram, 1967, p. 39).

Einstein's opposition to OQT—arising from the lack of realism of the theory—was implacable, even vehement. It was this, after all, which had led to the great rupture between mainstream theoretical physics and Einstein's own work. From 1905 to 1926 Einstein was at the centre of developments in theoretical physics. But from 1926 onwards the ways parted, essentially because Einstein was not able to bring himself to contribute to the development of OQT (confining himself to critical analysis of it). Robert Shankland, who met Einstein a number of times during the years 1950-4, has remarked on the uncharacteristic vehemence of Einstein's opposition to OQT.

His well-known scepticism on this subject [of quantum mechanics] was clearly evident and his comments on both the subject itself and its leading proponents were often highly critical and even emotional, in contrast to his restrained and quiet explanations of relativity (French 1979, p. 39).

Something of the strength of Einstein's opposition to OQT also emerges from a correspondence which he had with Born on the subject. Einstein makes it quite clear that he finds OQT unacceptable because of its lack of realism. Born persists in a stance of somewhat patronizing incomprehension, Einstein rather sharply writes that he does not wish to continue the discussion, and Pauli is obliged to step in and tick Born off for misunderstanding Einstein, even though he agrees with Born that Einstein's position

amounts to asking 'how many angels are able to sit on the point of a needle' (see Born, 1971, pp. 199-229).

The strongest statement of Einstein against OQT that I have come across is quoted by Fine (1986, p. 1): 'This theory [the present quantum theory] reminds me a little of the system of delusions of an exceedingly intelligent paranoiac, concocted of incoherent elements of thoughts.'

These quotations establish beyond all possible doubt that Einstein was committed to full-blooded scientific realism, at least as far as the basic aim of physics is concerned.

Einstein is absolutely correct to hold that OQT cannot be interpreted realistically. As he points out in his letter to Schrödinger, if one attempts to interpret the ψ -function of OQT as providing a complete description of reality, one is led to the (apparently) absurd conclusion that Schrödinger's cat persists as a superposition of being alive and being dead until we open the box and look. And similarly, we would have to conclude that the outcome of any quantum measurement is not some definite state of the apparatus but rather a superposition of macroscopically distinct states—the superposition only collapsing miraculously when we look. The simplest way to demonstrate the impossibility of interpreting OQT realistically, however, arises from the following consideration. If we interpret the ψ -function as describing quantum reality directly, and exclude *measurement* from the basic postulates of the theory, we are left with a theory that is fully deterministic, since quantum states, corresponding to ψ -functions, evolve deterministically in accordance with Schrödinger's equation. Such a version of OQT fails to make contact with the most basic feature of the quantum world—its probabilistic character. In short, just as Einstein declares, OQT must be regarded as a theory which makes probabilistic predictions about the results of performing *measurements* on systems, but which does not specify the actual physical state of the individual system in the absence of measurement.

How did this extraordinary state of affairs arise? Essentially because, as quantum theory (QT) developed with the work of Bohr, Heisenberg, Schrödinger, Born and others, no solution was

found to the quantum wave/particle problem. As we have seen, this problem was first discovered by Einstein with his invention of light quanta—or 'photons' as they subsequently came to be called. The problem was further intensified in 1923 when de Broglie proposed that *electrons*, up till then believed to be particles, have a wave-like aspect associated with them, as was subsequently confirmed experimentally by Davisson and Germer. In order to develop QT as a realistic theory, it would have been necessary to solve the quantum wave/particle problem in such a way that it is possible to specify, consistently and precisely, what sort of physical entities photons and electrons are as they evolve in space and time independently of measurement. This did not happen. Instead, Heisenberg invented matrix mechanics in 1925 intending, from the outset, that the theory should predict the outcome of measurements but should remain silent about what exists physically in the absence of measurement. Schrödinger invented wave mechanics in 1926 with the hope that the wave aspect of quantum entities would turn out to be fundamental. This hope was dashed when it became clear that the ψ -function could not be regarded as describing quantum reality directly, but had to be interpreted as containing probabilistic information about the results of performing *measurements* on an ensemble of similarly prepared systems—as Born was the first to point out.

We can begin to see some of the reasons for Einstein's vehement rejection of OQT as a satisfactory theory (despite its immense empirical success). It was Einstein after all who, in a sense, invented quantum theory. Planck introduced the idea that the energy E of an oscillator of frequency ν is quantized in accordance with $E = nh\nu$ as a calculational device, not as a new hypothesis incompatible with classical physics. Planck's aim was to deduce his empirical law of blackbody radiation from the basic postulates of classical physics. He was dismayed to discover that the quantization of energy *contradicted* classical physics, and he spent the next fifteen years or so trying to remove this defect from his derivation. It was Einstein, and Einstein alone, who appreciated that Planck's work spelt the downfall of classical physics, a new beginning being required. In this sense, Einstein initiated quantum

theory with his paradoxical 'heuristic' hypothesis that light consists of discrete quanta with energy $E = h\nu$, even though light also undeniably has a continuous wave-like character. For Einstein around 1905, the fundamental task of the new theory, needed to replace classical physics, would be to solve the riddle of the nature of quantum reality in view of its ostensibly contradictory particle and field aspects. No wonder Einstein was dismayed when the new theory was developed deliberately to *evade* and not to *solve* this basic quantum riddle.

But there is more than this to Einstein's opposition to OQT. As I have stressed above, the failure to solve the quantum wave/particle problem ensures that OQT cannot be interpreted realistically, which in turn ensures that OQT must be interpreted as making (probabilistic) predictions about the results of performing measurements. But this in turn has a variety of disastrous— though rarely noticed—consequences. For it means that OQT only issues in actual physical predictions if some part of *classical physics* (CP) is adjoined to OQT for a treatment of measurement. OQT alone can only issue in *conditional* predictions of the type: *if* a measurement of observable A is made, the outcome will be one or other of the values ($a_1 \dots a_n$) with probabilities ($p_1 \dots p_n$), with:

$$\sum_{r=1}^n p_r = 1.$$
 And even this goes too far: strictly speaking, according

to OQT, a quantum mechanical state ψ can only be attributed to a system in so far as the system has been subjected to some preparation procedure—which must be specified by means of CP. Thus OQT, devoid of CP, has no physical content whatsoever. It is only OQT + CP which has physical content. But OQT + CP, considered as a fundamental theory of physics, is a disaster. It is (i) grossly *ad hoc* or aberrant, in that it consists of two conceptually incoherent parts, OQT and CP. It is (ii) *imprecise*, because the circumstances in which CP is to be applied are only specified in terms of measurement, and the notion of measurement cannot be made precise (Maxwell, 1972b). It is (iii) *ambiguous* because the theory does not decide unambiguously between probabilism and determinism. It is (iv) *non-explanatory*, not only because of the *ad*

hoc character of the theory, but also because the theory is obliged to presuppose some part of what it is intended to explain, namely CP. The theory is (v) *severely restricted in scope* in that it cannot be applied to conditions which exclude the possibility of measurement, such as early states of the universe. It (vi) *excludes the possibility of quantum gravity and quantum cosmology*, since these would require measuring instruments, described in terms of CP, to exist outside space-time and beyond the cosmos, and clearly this is not possible. (These are points I have developed over a number of years: see Maxwell (1972b, 1973a, 1973b, 1975, 1976b, 1982), and especially (1988), pp. 1-8.)¹⁵³

These six gross defects—especially (i) to (iv)—ensure that OQT + CP is unacceptable as a fundamental physical theory. OQT + CP cannot justifiably be held to be part of theoretical scientific *knowledge*. (OQT + CP encompasses a great deal of *empirical* knowledge, but cannot be said to be an acceptable theory, constituting *theoretical* knowledge.) OQT + CP is as unacceptable as the absurd, empirically successful but grossly *aberrant* theories considered in previous chapters. In practice this point is beyond dispute. The vast majority of physicists, from soon after 1926 down to the present day (or down to 1993, when this chapter was first published), have regarded OQT as an entirely acceptable part of scientific knowledge: they have been able to do this because they have been able to pretend that OQT + CP is really just OQT. In almost all the textbooks and physical journals quantum theory is treated *as if* its postulates are purely quantum mechanical ones. As a result, OQT appears to be thoroughly *non-ad hoc*, precise and explanatory, as conceptually coherent and unified as any classical theory. But all this is an illusion. It is the outcome of pretending that the *physical* theory—the theory that has *physical content*—is OQT rather than OQT + CP. No such thing is possible. OQT, devoid of CP, has no physical content whatsoever. Only an all-pervasive intellectual dishonesty makes it possible to pretend that OQT alone has physical content (or that OQT + CP is really, somehow, just OQT).

¹⁵³ See also my [1994; 1995; 1998; 2004; 2010].

All this demonstrates just how sound Einstein's instincts were when he judged OQT to be an unacceptable theory. How unerringly correct Einstein was to declare that Bohr and company 'do not see what sort of risky game they are playing with reality'; and how sound his comparison is between OQT and 'the system of delusions of an exceedingly intelligent paranoiac, concocted of incoherent elements of thoughts'—namely QT *and* CP!

It is important to appreciate that the above six defects of OQT, even though consequences of the impossibility of interpreting OQT realistically, are not defects which only realists will recognize. Any physicist, whether realist or instrumentalist, aim-oriented empiricist or standard empiricist ought in practice, to regard the above defects sufficient grounds for finding OQT unacceptable. We have here, in effect, an additional general argument against instrumentalism and for realism. Any fundamental physical theory, and not just OQT, which is interpreted instrumentalistically as predicting only the (observable) outcomes of measurements will be, in the same way, unacceptably (i) *ad hoc*, (ii) imprecise, and (iv) non-explanatory. In other words, theoretical unity implies realism; anti-realism, built into a physical theory (as it is built into OQT) must inevitably, at some point, lead to unacceptable *ad hocness* or aberrance (see Maxwell, 1993b).

Even though it is not essential to be an aim-oriented empiricist in order to find OQT unacceptable, it helps. For aim-oriented empiricism provides a clear and cogent *raison d'etre* for finding OQT unacceptable even though the theory has met with such outstanding empirical success. Standard empiricism, on the other hand, can provide no such *raison d'etre*. If scientific theories ought in the end to be judged solely on the basis of empirical success and failure, then there can be no rational grounds for rejecting OQT, given its immense, its unprecedented empirical success.

It is in just this way that most of Einstein's contemporaries tended to view his rejection of OQT: as the outcome of unscientific, metaphysical prejudice, or even as an indication of 'senility' (as Einstein himself put it). Even Abraham Pais, so knowledgeable about, and so sympathetic towards Einstein, nevertheless regards Einstein's objections to OQT as 'unfounded' (Pais, 1982, p. 464).

Einstein's attitude towards OQT, so strikingly at odds with most of his contemporaries, provides further evidence in support of my contention that aim-oriented empiricism is implicit in Einstein's scientific work. If Einstein had assessed OQT in purely standard empiricist terms, he could have had no rational grounds for rejecting OQT—no grounds even for rejecting OQT as a 'point of departure' (since this is an *epistemological* judgement, to the effect that OQT is fundamentally *false*). From the standpoint of standard empiricism, Einstein's implacable opposition to OQT is just plain irrational prejudice. From the standpoint of aim-oriented empiricism, however, Einstein's rejection of OQT emerges as entirely well-founded, scientific, rational and objective. OQT is entirely acceptable from the standpoint of empirical considerations, but unacceptable from the equally important standpoint of theoretical unity, comprehensibility. The scandal is that the majority of contemporary physicists do not see this obvious point.

What is irrational, in other words, is not Einstein's rejection of OQT, but the majority *acceptance* of OQT, the general *blindness* to its gross defects. Einstein, I believe, held this to be the result of the fact that too many physicists put fame before understanding the universe. Einstein felt that, given a choice between winning a Nobel prize and improving our understanding of the universe, too many physicists would choose the former over the latter. This, for Einstein, amounted to a betrayal of the soul of theoretical physics, the pursuit of a corrupt goal, fame (not for Einstein so very different from the pursuit of power), in preference to the pursuit of the noble goal of improving understanding. And this in turn was, for Einstein, I believe, characteristic of a general sickness of our age: the pursuit of shallow or corrupt goals in life in preference to goals of genuine value.¹⁵⁴ Here is the source of Einstein's inability

¹⁵⁴ 'Perfection of means and confusion of goals seem—in my opinion—to characterize our age' (Einstein [1973], p. 337). One can regard this state of affairs as the result of the failure of our age to develop and implement a kind of rational inquiry designed to help us improve our goals, informed by aim-oriented rationalism and the philosophy of wisdom, themselves the outcome of generalizing Einstein's way of doing physics: see Maxwell [1976b and 1984].

to contribute to OQT after 1926. It is in this sense that Einstein's pursuit of his unified field theory is a kind of moral protest; this was the clearest way in which he could express his conviction as to what physics ought to be, at its best.

Einstein did not get everything right about OQT. He assumed that the ostensibly highly non-local features of OQT—which seem to contradict special relativity—do not correspond to reality. Here he was wrong.

If two particles, 1 and 2, interact at time t_1 and then separate widely then, in certain circumstances, a measurement performed on 1 at time t_2 enables one to predict with certainty what the result would be of measuring 2. A measurement of the momentum of 1 enables one to predict the precise momentum of 2; or, alternatively, a measurement of the position of 1 enables one to predict the position of 2. It is possible that 2 only acquires a precise momentum or position at time t_2 , when one or other kind of measurement is performed on 1. This possibility requires that an influence of some kind travels instantaneously from 1 to 2 to inform 2 as to whether it should acquire a precise momentum or position. If we reject the existence of such instantaneous influences, then in order to explain the correlations between measurements on 1 and 2 we are obliged to hold, it seems, that these correlations are the outcome of correlations established at time t_1 , when 1 and 2 interact. But this has the consequence that at time t_2 particle 2 must simultaneously have a precise position *and* momentum (since 2, by hypothesis, cannot 'know' instantaneously, at time t_2 , whether particle 1, far away, is subjected to a position or momentum measurement). But, according to OQT, no system can be in a state which corresponds to having simultaneously a precise position *and* momentum. Thus OQT implies that correlations cannot be established at t_1 when 1 and 2 interact; they must be established instantaneously, at t_2 , when one or other measurement is performed on 1.

That OQT does have this highly non-local character was discovered by Einstein, and was expounded in a famous paper by Einstein, Podolsky and Rosen (1935). Because of the evident clash with special relativity, Einstein concluded that this kind of non-local prediction of OQT is *false*. Particle 2 *does* have a precise

position and momentum at time t_2 irrespective of whether measurements are performed on 1 or not, and QT must be interpreted as a purely statistical theory which gives only an *incomplete* description of the evolution of the individual system.

Einstein held that the only reasonable option available was to interpret QT in this way, as an inherently incomplete, statistical theory of 'particles'. There can be no doubt that this reinforced his conviction that QT did not constitute a proper starting point for future developments—which in turn reinforced Einstein's search for a unified field theory.

Subsequent developments, due to Bohm (1957), Bell (1964), Aspect *et al.* (1982) and others, have shown that Einstein was *wrong* to dismiss the nonlocal predictions of OQT as not corresponding to reality: these predictions have now been experimentally confirmed!

This concludes my case for saying Einstein invented and applied aim-oriented empiricism in scientific practice in developing the special and general theories of relativity, and in critically examining quantum theory.

It may be asked: but did Einstein explicitly *advocate* aim-oriented empiricism? I turn now to a discussion of this question.

There can be, to begin with, no doubt that Einstein devoted his life to the goal of discovering the unified structure of the universe and that, for him, this constituted an entirely proper aim for science, indeed the noblest motive for pursuing scientific inquiry. Something of what the desire to understand meant to Einstein emerges from the following passage.

The most beautiful experience we can have is the mysterious. It is the fundamental emotion which stands at the cradle of true art and true science. Whoever does not know it and can no longer wonder, no longer marvel, is as good as dead, and his eyes are dimmed. It was the experience of mystery—even if mixed with fear—that engendered religion. A knowledge of the existence of something we cannot penetrate, our perceptions of the profoundest reason and the most radiant beauty, which only in their most primitive forms are accessible to our minds—it is this knowledge and this emotion that

constitute true religiosity; in this sense, and in this alone, I am a deeply religious man. I cannot conceive of a God who rewards and punishes his creatures, or has a will of the kind that we experience in ourselves. Neither can I nor would I want to conceive of an individual that survives his physical death; let feeble souls, from fear or absurd egoism, cherish such thoughts. I am satisfied with the mystery of the eternity of life and with the awareness and a glimpse of the marvelous structure of the existing world, together with the devoted striving to comprehend a portion, be it ever so tiny, of the Reason that manifests itself in nature (Einstein (1973), p. 11).

On one occasion in 1925 he expressed himself to the novelist Esther Salaman in the following terms:

I want to know how God created this world. I'm not interested in this-or-that phenomenon, in the spectrum of this-or-that element. I want to know His thoughts, the rest are details. (Salaman (1979), p. 22.)

That a basic aim of science is to unify all phenomena is affirmed in numerous passages, such as, from 1936 (see Einstein, 1973, p. 293):

The aim of science is, on the one hand, a comprehension, as *complete* as possible, of the connection between the sense experiences in their totality, and, on the other hand the accomplishment of this aim *by use of a minimum of primary concepts and relations*. (Seeking, as far as possible, logical unity in the world picture, i.e. paucity in logical elements.)

Einstein also makes it clear that science at its best assumes that this goal of unification is realizable.

Certain it is that a conviction, akin to religious feeling, of the rationality or intelligibility of the world lies behind all scientific work of a higher order. This firm belief, a belief bound up with deep feeling in a superior mind that reveals itself in the world of experience, represents my conception of God. In common parlance this may be described as 'pantheistic' (Spinoza). (First published 1929; see Einstein (1973, p. 262.)

And on another occasion:

From the very beginning there has always been present the attempt to find a unifying theoretical basis for all [the] single sciences, consisting of a minimum of concepts and fundamental relationships, from which all the concepts and relationships of the single disciplines might be derived by logical process. This is what we mean by the search for a foundation of the whole of physics. The confident belief that this ultimate goal may be reached is the chief source of the passionate devotion which has always animated the researcher. (First published 1940; see Einstein, 1973, p. 324.)

As for scientific realism, Einstein expresses himself with his usual clarity and brevity:

The belief in an external world independent of the perceiving subject is the basis of all natural science. Since, however, sense perception only gives information of this external world or of "physical reality" indirectly, we can only grasp the latter by speculative means. (First published 1931; see Einstein, 1973, p. 266.)

And, on another occasion, as we have already seen:

Physics is an attempt to grasp reality as it is thought independently of its being observed.

It is all summed up succinctly in a letter to Cornelius Lanczos in 1942:

You are the only person I know who has the same attitude toward physics as I have: belief in the comprehension of reality through something basically simple and unified. (Dukas and Hoffmann, 1979, p. 68.)

All this might seem more than enough to demolish decisively the views of those, like Fine and Popper, who hold that Einstein upheld some version of standard empiricism. Unfortunately it is not. In all the above quotations, Einstein can be interpreted as asserting no more than that he, and science, seek to discover, and presuppose the existence of, a unified structure to the universe, *in the context of discovery*. According to this interpretation, Einstein would hold that, *in the context of justification*, nothing must be permanently assumed about the nature of the universe, the sole aim

being empirical adequacy, empirical considerations *alone* in the end deciding what is to constitute theoretical scientific knowledge.

On this issue—the crucial issue which divides off standard from aim-oriented empiricism—Einstein seems to have wavered. Consider the following passage.

Can we hope to be guided safely by experience at all when there exist theories (such as classical mechanics) which to a large extent do justice to experience, without getting to the root of the matter? I answer without hesitation that there is, in my opinion, a right way, and that we are capable of finding it. Our experience hitherto justifies us in believing that nature is the realization of the simplest conceivable mathematical ideas. I am convinced that we can discover by means of purely mathematical constructions the concepts and the laws connecting them with each other, which furnish the key to the understanding of natural phenomena. Experience may suggest the appropriate mathematical concepts, but they most certainly cannot be deduced from it. *Experience remains, of course, the sole criterion of the physical utility of a mathematical construction.* But the creative principle resides in mathematics. In a certain sense, therefore, I hold it true that pure thought can grasp reality, as the ancients dreamed. (Italics mine.) (First published 1933; see Einstein, 1973, p. 274.)

This comes tantalizingly close to aim-oriented empiricism. A central tenet of aim-oriented empiricism is that we are rationally entitled to assume that the universe is knowable—there being some fallible, non-mechanical but rational method of discovery available to us—the knowability of the universe implying its comprehensibility. It is just this key element of aim-oriented empiricism which Einstein asserts here, his epistemological and methodological instincts as usual getting almost everything right. (The point is also brilliantly made in one of Einstein's most famous sayings: 'Raffiniert ist der Herrgott, aber boshaft ist er nicht'—God is sublime but not malicious.) Unfortunately, in the quotation given above, the italicized sentence provides Popper or Fine with the perfect excuse for interpreting the passage as a defence of standard

empiricism. One can argue, of course, that the whole passage only really makes sense if interpreted as asserting: experience remains the *sole* criterion of physical utility *granted that we restrict our attention to simple, unified theories*. This would of course *violate* standard empiricism. But the text, as it stands, is sufficiently ambiguous to leave the matter undecided. Consider next the following passage.

The very fact that the totality of our sense experiences is such that by means of thinking (operations with concepts, and the creation and use of definite functional relations between them, and the coordination of sense experiences to these concepts) it can be put in order, this fact is one which leaves us in awe, but which we shall never understand. One may say 'the eternal mystery of the world is its comprehensibility'. It is one of the great realizations of Immanuel Kant that the postulation of a real external world would be senseless without this comprehensibility. In speaking here of 'comprehensibility', the expression is used in its most modest sense. It implies: the production of some sort of order among sense impressions, this order being produced by the creation of general concepts, relations between these concepts, and by definite relations of some kind between the concepts and sense experience. It is in this sense that the world of our sense experiences is comprehensible. The fact that it is comprehensible is a miracle.

In my opinion, nothing can be said *a priori* concerning the manner in which the concepts are to be formed and connected, and how we are to coordinate them to sense experiences. In guiding us in the creation of such an order of sense experiences, success alone is the determining factor. All that is necessary is to fix a set of rules, since without such rules the acquisition of knowledge in the desired sense would be impossible. One may compare these rules with the rules of a game in which, while the rules themselves are arbitrary, it is their rigidity alone which makes the game possible. However, the fixation will never be final. It will have validity only for a special

field of application (i.e., there are no final categories in the sense of Kant). (First published 1936; see Einstein 1973, p. 292.)

This, once again, is tantalizingly close to aim-oriented empiricism. Einstein recognizes clearly that only in a very special kind of universe—a comprehensible universe—is scientific explanation and understanding possible. He recognizes that the particular way the universe is assumed to be comprehensible at any stage in the development of science will lead to rules or principles—such as Galilean or Lorentz invariance, the principle of equivalence, conservation of momentum and energy—without which physics would be impossible. And he points out that these rules are not final: it is to be expected that they will change as science advances. All this accords beautifully with aim-oriented empiricism. What violates aim-oriented empiricism is the suggestion that there is no Kantian synthetic *a priori* proposition built into scientific knowledge. According to aim-oriented empiricism, there is just one such proposition, namely: the universe is comprehensible (in some way or other). We cannot of course *know for certain* that this proposition is true. It must remain for ever a *conjecture*—all our knowledge being conjectural in character. The decisive point is that knowledge becomes impossible if this conjecture is false, this rationally entitling us to adopt the conjecture as a permanent part of scientific knowledge. Nothing is to be gained from doubting the conjecture, and much may be gained from incorporating it into the rest of our conjectural scientific knowledge. In short, 'the universe is comprehensible' is a synthetic *a priori* statement not in the full-blooded Kantian sense that it can be known to be true of all possible experience with absolute certainty, but in the radically qualified Kantian sense that it is a conjecture about reality—about the noumenal world—which must remain permanently an integral part of conjectural human knowledge, and which is adopted as knowledge on non-empirical grounds. This crucial tenet of aim-oriented empiricism is, it seems, explicitly rejected by Einstein in the above passage—even though the whole point of the passage, ironically enough, is to affirm it, affirm, that is, that science *cannot* proceed without the assumption that the universe is comprehensible.

Einstein's ambivalent attitude to the crucial issue which separates off standard from aim-oriented empiricism gains explicit expression in the following quotation.

[The aim of science is to arrive] at a system of the greatest conceivable unity, and of the greatest poverty of concepts of the logical foundations, which is still compatible with the observations made by our senses. We do not know whether or not this ambition will ever result in a definite system. If one is asked for his opinion, he is inclined to answer no. While wrestling with the problems, however, one will never give up the hope that this greatest of all aims can really be attained to a very high degree. (First published 1936; see Einstein, 1973, p. 294.)

We might interpret this to mean that when Einstein is thinking primarily as a theoretical physicist he unthinkingly takes the ultimate comprehensibility of the universe for granted—the key component of aim-oriented empiricism. When he comes to reflect philosophically about the aims and methods of his work, however, his (misconceived) philosophical conscience gets the better of him, and he lapses into standard empiricism. Einstein's scientific instincts, in short, are more enlightened than his philosophical reflections—an important point, implicit in my claim that aim-oriented empiricism arose, for Einstein, out of scientific practice, adopted in response to a severe scientific problem.

Are we to conclude, then, that Einstein did not in the end manage to free himself explicitly from the trap of standard empiricism? One point to remember is that throughout his scientific life Einstein's views on the philosophy of science evolved from something close to Machian positivism at the outset (an extreme version of standard empiricism) to a view that comes to resemble aim-oriented empiricism more and more closely towards the end of his life. Einstein himself put the matter like this, in a letter to Lanczos in 1938.

Coming from sceptical empiricism of somewhat the kind of Mach's, I was made, by the problem of gravitation, into a believing rationalist, that is, one who seeks the only trustworthy source of truth in mathematical simplicity. The logically simple does not, of course, have to be

physically true; but the physically true is logically simple, that is, it has unity at the foundation (see Holton, 1973, p. 241).

This, to begin with, sounds like a clear enough confession of a convinced aim-oriented empiricist. As it stands, it is perhaps something of an oversimplification. In the first place, as we have seen above, elements of aim-oriented empiricist thinking can be found in Einstein's scientific work almost from the outset—from Einstein's first great creative period in 1902-5. Second, Einstein's views concerning the philosophy of science went on developing long after the creation of general relativity, right to the end of his life. Our best hope, then, of finding a clear, unambiguous formulation of aim-oriented empiricism is to look at Einstein's very last writings on philosophy of science. I provide two final quotations. The first comes from Einstein's 'Autobiographical Notes', written when he was 67. Einstein is discussing the points of view from which physical theories can be critically assessed, quite generally.

The first point of view is obvious: the theory must not contradict empirical facts. . . . The second point of view is not concerned with the relation to the material of observation but with the premises of the theory itself, with what may briefly but vaguely be characterized as the 'naturalness' or 'logical simplicity' of the premises (of the basic concepts and of the relations between these which are taken as a basis). This point of view, an exact formulation of which meets with great difficulties, has played an important role in the selection and evaluation of theories since time immemorial. The problem here is not simply one of a kind of enumeration of the logically independent premises (if anything like this were at all unequivocally possible), but that of a kind of reciprocal weighing of incommensurable qualities ... Of the 'realm' of theories I need not speak here, inasmuch as we are confining ourselves to such theories whose object is the *totality* of all physical appearances. The second point of view may briefly be characterized as concerning itself with the 'inner perfection' of the theory, whereas the first

point of view refers to the 'external confirmation'. The following I reckon as also belonging to the 'inner perfection' of a theory: We prize a theory more highly if, from the logical standpoint, it is not the result of an arbitrary choice among theories which, among themselves, are of equal value and analogously constructed (Einstein, 1969, pp. 21-3).

It is surely clear from this that Einstein came quite explicitly to repudiate all versions of standard empiricism towards the end of his life. There is no suggestion here that the second requirement of 'inner perfection' or unity is somehow to be reduced to the first requirement of empirical adequacy: empirical considerations do *not*, for Einstein, *alone* determine choice of theory. Furthermore, Einstein has made it abundantly clear already that, in his view, in choosing only theories which satisfy the requirement of inner perfection, we are in effect assuming that the universe itself is comprehensible—this being a permanent presupposition of scientific knowledge upheld on non-empirical grounds. But in case there is any doubt on this score, here is a passage written in 1950 in which the thesis that there can be no knowledge without the presupposition that the universe is comprehensible is explicitly affirmed.

It is of the very essence of our striving for understanding that, on the one hand, it attempts to encompass the great and complex variety of man's experience, and that on the other, it looks for simplicity and economy in the basic assumptions. The belief that these two objectives can exist side by side is, in view of the primitive state of our scientific knowledge, a matter of faith. Without such faith I could not have a strong and unshakable conviction about the independent value of knowledge (Einstein, 1973, p. 357).

I conclude that Einstein came close to articulating aim-oriented empiricism towards the end of his life, even if he did not recognize that this position is required to solve the problem of induction, and did not appreciate that it provides a more rational conception of science than does standard empiricism—and not a less rational

conception, as Einstein's references to 'faith' and 'miracle-creed' tend to suggest.

In the end, however, what really matters is the philosophy of science implicit in Einstein's scientific deeds. Einstein himself held this view. As he put it: 'If you want to find out anything from the theoretical physicists about the methods they use, I advise you to stick closely to one principle: don't listen to their words, fix your attention on their deeds' (Einstein, 1973, p. 270). As we have seen above, in order to make rational sense of Einstein's scientific judgements and deeds it is essential to see them from the standpoint of aim-oriented empiricism. More important, Einstein can be said to have invented aim-oriented empiricism in scientific practice during the course of discovering the special and general theories of relativity. His success in discovering these theories owes much to the invention and exploitation of the rational method of discovery of aim-oriented empiricism. This aspect of Einstein's work transformed the whole character of subsequent theoretical physics. Einstein's contributions to theoretical physics are intimately interrelated with his contribution to the philosophy of physics: after Einstein, indeed, physics and philosophy of physics ought to form one integrated discipline—aim-oriented empiricist natural philosophy. The various versions of standard empiricism defended by Popper, van Fraassen and Fine (and most contemporary philosophers of science) all fail to do justice to this vital dimension of Einstein's contribution to science. Indeed, advocacy of standard empiricism after Einstein amounts in itself to a failure to understand an important aspect of Einstein's contribution to science.¹⁵⁵

¹⁵⁵ In my view, the most important implications of the new way of doing physics created by Einstein in developing special and general relativity lie in fields far beyond that of theoretical physics: see chs. 3, 7 and 8, and my (1976a, 1984 or 2007, 1986, 1991, 1992, and 2010a) where I attempt to spell out these implications for science as a whole, for technological research, social inquiry, scholarship and education.

Chapter Seven

The Enlightenment Programme and Karl Popper

(First published, in a shortened form, in *Karl Popper: A Centenary Assessment. Volume 1: Life and Times, Values in a World of Facts*, ed. I. Jarview, K. Milford and D. Miller, chapter 11, Ashgate, London, pp. 177-190.)

Abstract

By the Enlightenment Programme I mean the idea of the 18th century French Enlightenment, of learning from scientific progress how to go about making social progress towards world enlightenment. Karl Popper's work can be interpreted as contributing to this Programme. This is apparent, especially, in the way he generalizes falsificationism to form critical rationalism, and in the link that this provides between *The Logic of Scientific Discovery* and *The Open Society and Its Enemies*. But Popper does not, to my knowledge, explicitly advocate the Enlightenment Programme. Judged from the standpoint of contributing to the

Programme, Popper's work is important but not the final word on the subject. To begin with, Popper's falsificationism needs to be improved to take into account that metaphysical theses concerning the comprehensibility and knowability of the universe are built into the aims and methods of science. This leads to a new conception of the progress-achieving methods of science which, when generalized, lead to a new conception of rationality, with new implications for social inquiry, and the political task of devoting reason to developing a more civilized world.

1 Karl Popper's Most Significant Contribution

Karl Popper's most significant contributions are contained in his first four books: *The Logic of Scientific Discovery*, *The Open Society and Its Enemies*, *The Poverty of Historicism* and *Conjectures and Refutations*.

It is important to appreciate the existence of a central backbone of argument running through these four books. In *The Logic of Scientific Discovery* (1959), Popper argues (as we all know) that all scientific knowledge is irredeemably conjectural in character, it being impossible to verify theories empirically. Science makes progress by proposing bold conjectures in response to problems, which are then subjected to sustained attempted empirical refutation. This falsificationist conception of scientific method is then generalized to form Popper's conception of (critical) rationality, a general methodology for solving problems or making progress. As Popper puts it in *The Logic of Scientific Discovery* "inter-subjective *testing* is merely a very important aspect of the more general idea of inter-subjective *criticism*, or in other words, of the idea of mutual rational control by critical discussion".¹⁵⁶

In order to make sense of the idea of *severe* testing in science, we need to see the experimentalist as having at least the germ of an idea for a rival theory up his sleeve (otherwise testing might degenerate into performing essentially the same experiment again and again). This means experiments are always *crucial experiments*, attempts at trying to decide between two competing

¹⁵⁶. See Popper (1959, p. 44, n *1). See also Popper (1963, pp. 193-200); (1972, pp. 119 & 243); (1976a, pp. 115-6).

theories. Theoretical pluralism is necessary for science to be genuinely empirical. And, more generally, in order to *criticize* an idea, one needs to have a rival idea in mind. Rationality, as construed by Popper, requires plurality of ideas, values, ways of life. Thus, for Popper, the rational society *is* the open society. Given pre-Popperian conceptions of reason, with their emphasis on proof rather than criticism (and associated plurality of ideas), the idea that the rational society is the open society is almost a contradiction in terms. There is thus a very close link between *The Logic of Scientific Discovery*, on the one hand, and *The Open Society and Its Enemies*, *The Poverty of Historicism* and *Conjectures and Refutations* on the other.

And the direction of argument does not go in just one direction, from *The Logic of Scientific Discovery* to *The Open Society and Its Enemies*: it goes in the other direction as well. For in *The Open Society and Its Enemies* Popper argues that rationality, and scientific rationality as well, need to be conceived of in social and institutional terms (and the argument is echoed in *The Poverty of Historicism*, in connection with a discussion about the conditions required for scientific progress to be possible).¹⁵⁷ *The Open Society and Its Enemies*, *The Poverty of Historicism* and *Conjectures and Refutations* illuminate and enrich the doctrines of *The Logic of Scientific Discovery*.

2 The Enlightenment Programme

Much of the importance of Popper's first four books stems from the fact that they constitute a major contribution to what may be called "The Enlightenment Programme" – the basic idea of the 18th century Enlightenment, especially the French Enlightenment, of *learning from scientific progress how to achieve social progress towards an enlightened world*. Popper's work does much to revitalize and improve on the version of the Enlightenment Programme that we have inherited from the 18th century Enlightenment, from Voltaire, Diderot, Condorcet and the other *philosophes*.¹⁵⁸ But, as we shall see – and this is the main point of

¹⁵⁷. Popper (1966a, 217-20); (1961, 154-7).

¹⁵⁸. The best account of the Enlightenment known to me is Gay (1973).

this talk – Popper's version of the Enlightenment Programme, despite its great virtues, is still defective, and needs further improvement.

But before I discuss Popper's contribution, I want first to say a few words about how profoundly important the basic Enlightenment idea is of learning from scientific progress how to achieve social progress towards an enlightened world.

Science has made astonishing progress in improving knowledge. But social progress towards an enlightened world seems much more problematic. This discrepancy would be reason enough to take very seriously indeed the Enlightenment idea of seeing whether we can learn from scientific progress how to achieve greater social progress towards an enlightened world. But what makes this Enlightenment idea so much more important and urgent, for our times, is that all too often, and tragically, modern science and technology have actually been implicated in some of our worst human disasters. Modern science and technology have undoubtedly done much to help relieve human suffering and enhance the quality of human life; but there is still our terrible record of unnecessary human suffering and death, our record of man-made disasters during the past hundred years or so: horrifyingly destructive wars; the terrifying threat posed by modern armaments, conventional, chemical, biological and nuclear; vast inequalities in wealth across the globe; explosive population growth; the destruction of natural habitats and the rapid extinction of species; pollution of earth, sea and air, the latter leading to the thinning of the ozone layer and to global warming. And a crucial point to note about these global problems is that they have been made possible, have even, in a perfectly legitimate sense, been *caused*,¹⁵⁹ by the advent of modern science and

¹⁵⁹. It may be objected: it is not *science* that is the cause of our global problems but rather the things that we *do*, made possible by science and technology. This is obviously correct. But it is also correct to say that scientific and technological progress *is* the cause. The meaning of "cause" is ambiguous. By "the cause" of event E we may mean something like "the most obvious observable events preceding E that figure in the common sense explanation for the occurrence of E". In this sense, human actions (made possible by science) are the cause of such things as people being killed in war, destruction of tropical rain forests. On the

technology. Without the amazing success of modern science and technology, they would not have happened. There is nothing surprising about this. New scientific knowledge and technological know-how enormously increase our power to act: in the absence of enlightenment, of wisdom, our new power to act will sometimes have good consequences, but will also, as often as not, have bad consequences, whether intended, as in war, or unintended, as in global warming. Before the advent of modern science, lack of wisdom, of enlightenment, did not matter too much; we lacked the power to do too much damage to ourselves and the planet. But now that we do have modern science, and the unprecedented powers that it has given us, lack of enlightenment puts us into a position of unprecedented peril. It may even be that our very survival depends on humanity learning a bit more wisdom. Instead of blaming science for our troubles, as many do, we need, rather, to see whether we can learn from the astonishing and dangerous success of science about how to acquire a bit more global wisdom.

The 18th century *philosophes* interpreted the basic Enlightenment idea as requiring that the social sciences be developed alongside the natural sciences. Francis Bacon had already argued that, in order to better the lot of humanity, it is essential to improve our knowledge of the natural world. The *philosophes*, understandably enough, came to the conclusion that, in order to make social progress it is, if anything, even more important to improve knowledge of the social world. So they set about creating and developing the social sciences: economics, psychology, history, anthropology, sociology, political science. This was continued throughout the 19th century by such men as

other hand, by the "cause" of E we may mean "that prior change in the environment of E which led to the occurrence of E, and without which E would not have occurred". If we put the 20th century into the context of human history, then it is entirely correct to say that, in this sense, scientific-and-technological progress is the cause of distinctively 20th century disasters: what has changed, what is new, is scientific knowledge, not human nature. Yet again, from the standpoint of theoretical physics, "the cause" of E might be interpreted to mean something like "the physical state of affairs prior to E, throughout a sufficiently large spatial region surrounding the place where E occurs". In this third sense, the sun continuing to shine is as much a part of the cause of war and pollution as human action or human science and technology.

Saint-Simon, Comte, Mill and Marx until, by the mid-20th century, Departments of these various social sciences, as conceived of by the 18th century, had been created in Universities all over the world.¹⁶⁰ It is hardly too much to say that academic inquiry, as it exists today, is the outcome of developing and institutionalizing the scientific revolution of the 17th century, and the Enlightenment Programme of the 18th century.

The Enlightenment was of course opposed by the Romantic movement. This put its faith in emotion, imagination, spontaneity, inspiration, art, genius, and opposed the Enlightenment faith in science and reason. And this Romantic opposition to the Enlightenment, to science and reason, is still influential today in such fields as politics, education, anti-science movements, the arts. And it is still influential in some parts of academic inquiry, in such areas as philosophy, cultural studies, anthropology. Postmodernism comes out of the Romantic movement.¹⁶¹

But in objecting to the rationalism of the Enlightenment, Romanticism entirely missed the point. For the Traditional Enlightenment, inherited from the 18th century, suffers, not from too much reason, but from not enough. It amounts to a characteristic kind of irrationality masquerading as rationality. In developing the basic Enlightenment idea intellectually, the *philosophes* botched the job; and unfortunately it is this botched, irrational version of the Enlightenment Programme that we now have built into the intellectual and institutional structure of the academic enterprise. Academic inquiry today, when judged from the standpoint of helping us create a better world, is damagingly irrational in a wholesale, structural way.¹⁶²

Three steps that need to be got right to put the basic Enlightenment idea into practice correctly:

(i) The progress-achieving methods of science need to be correctly identified.

¹⁶⁰. See Hayek (1979); Fargaus (1993, Introduction).

¹⁶¹. For accounts of Romantic opposition to the Enlightenment see: Berlin (1980); (1999); Gascardi (1999).

¹⁶². I have argued for this theses in much greater detail in Maxwell (1984a).

(ii) These methods need to be correctly generalized so that they become fruitfully applicable to any worthwhile, problematic human endeavour, whatever the aims may be, and not just applicable to the endeavour of improving knowledge.

(iii) The correctly generalized progress-achieving methods then need to be exploited correctly in the great human endeavour of trying to make social progress towards an enlightened, civilized world.

Unfortunately, the *philosophes* of the Enlightenment got all three points disastrously wrong. They failed to capture correctly the progress-achieving methods of natural science (in that they defended inductivist, or at least verificationist, conceptions of science); they failed to generalize these methods properly; and, most disastrously of all, they failed to apply them properly so that humanity might learn how to become more civilized or enlightened by rational means. Instead of applying the generalized progress-achieving methods of science to *social life itself*, so that social progress might be achieved, the *philosophes* sought to apply scientific method merely to *social science*. Reason (as construed by the *philosophes*) got applied, not to the task of making social progress towards an enlightened world, but to the task of making intellectual progress towards greater knowledge about the social world. Social inquiry was developed, not as social methodology or social philosophy, but as social science.

That the *philosophes* made these blunders in the 18th century is forgivable; what is unforgivable is that these blunders still remain unrecognized and uncorrected today, over two centuries later. Instead of correcting the blunders, we have allowed our institutions of learning to be shaped by them as they have developed throughout the 19th and 20th centuries, so that now the blunders are an all-pervasive feature of academia, as we shall see in more detail in a moment.

3 Popper's Contribution to the Enlightenment Programme

As I have already indicated, Popper made enormously important improvements to the Enlightenment Programme inherited from the 18th century. Inductivism and verificationism become falsificationism. Traditional conceptions of reason, with all the

emphasis on proof and justification, become critical rationalism. These improvements, at steps (i) and (ii) of the Enlightenment Programme become all-important when it comes to step (iii): the application of reason to social life, to politics, to problems of living, to political philosophy. As a result of bringing about a revolution in our conception of scientific method, and of rationality more generally, Popper in effect transforms the very idea of "the rational society", so that this ceases to be something that is morally and politically abhorrent, and becomes both highly desirable, and achievable, instead.

Given traditional, pre-Popperian conceptions of science and reason, which tend to see science as establishing secure knowledge of truth by means of evidence, and tend to see reason as establishing truth by means of deductive argument,

the "rational society" can only be a society determined, or at least severely constrained, by "the rules of reason". Reason becomes a kind of tyrant. Individual liberty, diversity of views and ways of life, wayward imagination, disagreement and protest would all be suppressed by the iron rule of reason and logic. Granted such verificationist, authoritarian conceptions of reason,¹⁶³ the "rational society" can only be regarded as a kind of nightmarish totalitarian state, the very opposite of democracy and liberalism.¹⁶⁴ No wonder the Romantics protested.

But Popper's revolutionary ideas about science and reason change all this dramatically. First, granted Popper's falsificationist conception of scientific method, imagination plays a crucial role in science. Imagination is needed to dream up new wild speculations, subsequently to be submitted to ferocious attempts at empirical refutation. Second, plurality of conflicting theories is absolutely essential for scientific progress, not only to increase the store of theories to be submitted to attempted refutation, but in order to ensure that theories are severely tested in the first place. As I have already mentioned, in order to make sense of the idea of severe

¹⁶³. Significant, in this connection, is Popper's detection of an authoritarian streak in even apparently anti-authoritarian traditional conceptions of empiricism and reason upheld by, for example Bacon and Descartes: see Popper (1963, Introduction).

¹⁶⁴. For a novelist's fantasy of such a "rational" society, see Zamyatin (1972).

testing, we need to see the experimentalist as having at least the germ of an idea for a rival theory up his sleeve (otherwise testing might degenerate into performing essentially the same experiment again and again). This means experiments are always *crucial* experiments, attempts at trying to decide between two competing theories. Theoretical pluralism is necessary for science to be genuinely empirical.¹⁶⁵

Both these points carry over when Popper's falsificationist conception of scientific method is generalized to form critical rationalism. Reason, quite generally, is at a loss without imagination. Imagination is required to dream up possible solutions to problems, which can then be submitted to severe criticism. Again, plurality of views is an essential ingredient of Popper's conception of reason. Criticism can only deliver a good idea as to how to solve a problem if there are a plurality of ideas to criticize in the first place. And merely in order to criticize an idea, one needs to have some kind of rival idea in mind, at least as a possibility.

Rationality, as construed by Popper, requires plurality of ideas, values, ways of life, the freedom to imagine, to criticize authority, dogma and received opinion. It demands sustained tolerance of diversity of views and ways of life, together with the existence of traditions of criticism, so that good ideas may be selected from a pool of not-so-good ideas. Reason, as Popper emphasizes, needs to be seen in social, political and institutional terms (Popper, 1966a, ch. 24; 1961, section 32). Thus, granted Popper's revolutionary conceptions of scientific method and reason, the "rational society" is not some kind of totalitarian society, but just the opposite, the "open society" – a society that tolerates doubt, diversity of views and ways of life, criticism, and sustains individual liberty, reasonableness, humanity, justice and democracy. Reason, instead of being the enemy of freedom, individuality, imagination, democracy and justice, becomes the friend of these things, indeed

¹⁶⁵. This point is especially emphasized and further developed by Feyerabend (1965). Popper too emphasizes that, in order to make sense of the idea of severe testing we need to appeal to crucial experiments: see, for example, Popper (1963, 112).

essential for their preservation and development. As Popper puts it in a stray remark tossed out during the course of developing the argument: "We have to learn the lesson that intellectual honesty is fundamental for everything we cherish" (Popper, 1966a, vol. 2, p. 59).

In *The Open Society and Its Enemies*, Popper depicts an epic struggle between those who have sought to help sustain and promote the open society (i.e. the rational society), and those who have opposed it. And he shows how even some of the greatest thinkers of the past have been beguiled by false ideas of science and reason into arguing for the closed society, above all Plato and Marx.

In these ways, the path Popper pursues, from his conjectural, falsificationist conception of science to its generalization to form critical rationalism, and its application to some of the most urgent and profound political and social problems of our times, represents an immensely valuable rediscovery and transformation of the 18th century Enlightenment Programme. Popper's contribution is important and profound; but it is nevertheless defective. It needs further improvement. (Popper's followers, in so far as they refuse to consider the need for further improvement, do Popper a great disservice – and humanity, of course, an even greater disservice. At the heart of Popper's thought there is the insight that scientific method, and reason, rightly understood, deliver, not certainty, but rather uncertain progress, improvement, development, growth.¹⁶⁶ We betray Popper's philosophy quite fundamentally if we do not take it as a set of proposals, suggestions and arguments urgently in need of further development.)

In what follows I indicate how Popper's version of the Enlightenment Programme needs further improvement in two stages, which I shall discuss under the headings "The Improved Popperian Enlightenment" and "The New Enlightenment".

4 The Improved Popperian Enlightenment

¹⁶⁶. See, for example, Popper (1959, Preface to the English Edition, 1958); Popper (1963, Ch. 8 and 10).

A basic inadequacy of Popper's version of the Enlightenment Programme is that it depicts social inquiry, in highly traditional terms, to be social science, indeed pro-naturalist social science, with methods akin to those of the natural sciences. Popper's criticisms of some traditional views associated with social inquiry – his criticisms of historicism, historicist social science, and Utopian social engineering – are excellent and decisive. But he does not carry through this criticism of traditional views far enough; he fails to correct the greatest blunder of the 18th century Enlightenment, namely the mistake of applying reason to social *science* rather than to social *life*!

The basic Enlightenment idea, after all, is to learn from scientific progress how to make social progress towards an enlightened world. Putting this idea into practice involves getting appropriately generalized progress-achieving methods of science *into social life itself*! It involves getting progress-achieving methods into our institutions and ways of life in addition to science, into government, industry, agriculture, commerce, international relations, the media, the arts, education. But in sharp contrast to all this, the Traditional Enlightenment has sought to apply generalized scientific method, not to social *life*, but merely to social *science*! Instead of helping humanity learn how to become more civilized by rational means, the Traditional Enlightenment has sought merely to help social scientists improve knowledge of social phenomena (this knowledge then being applied to help solve social problems). The outcome is that today academic inquiry devotes itself to acquiring knowledge of natural and social phenomena, but does not attempt to help humanity learn how to become more civilized. Instead of social inquiry having, as its basic task, to promote cooperatively rational tackling of problems of living in the social world, its primary task, rather, is to acquire knowledge of social phenomena. Instead of being social methodology or social philosophy, social inquiry is pursued as social *science*.

This is the blunder that Popper simply reproduces, and fails to correct. Popper, one might almost say, argues for the open society, but fails to argue for open social inquiry, for a kind of open inquiry devoted to promoting the open society by rational means.

In order to correct this third, monumental and disastrous blunder, we need, as a first step, to bring about a revolution in the nature of academic inquiry, beginning with social inquiry and the humanities. Social inquiry needs to be, not social *science*, but rather social *methodology* or social *philosophy*, concerned to promote rational tackling of problems of living in the social world.¹⁶⁷

Let us now see, in a little more detail, what would result from correcting this third, monumental blunder of the Traditional Enlightenment. What we need to do is to see what results from applying the progress-achieving rules of reason (arrived at by generalizing the progress-achieving methods of science) to social *life* rather than to social *science*, to the task of making *social progress* towards a civilized world rather than to the task of making *intellectual progress* towards better knowledge of social phenomena.

In order to make clear what is at stake here, I need to appeal to a slightly improved version of Popper's critical rationalism, which I shall call "problem-solving rationality". This can be summed up like this:

1. Articulate, and try to improve the articulation of, the problem to be solved.
2. Propose and critically assess possible solutions.
3. When necessary, break up the basic problem to be solved into a number of preliminary, simpler, analogous, subordinate or specialized problems (to be tackled in accordance with rules 1 and 2, in an attempt to work gradually towards a solution to the basic problem to be solved.
4. Interconnect attempts to solve basic and specialized problems, so that basic problem-solving may guide, and be guided by, specialized problem-solving.¹⁶⁸

Popper's critical rationalism consists of rules 1 and 2; problem-solving rationality improves on this by adding on rules 3 and 4,

¹⁶⁷. Elsewhere I have spelled out in more detail than I am able to do here why this revolution is needed, and what it would amount to: see Maxwell (1984a). See also ch. 3 and my (1976a), (1991), (1992a), (1997b), (2001, ch. 9) and (2002a).

¹⁶⁸. For a version of these rules of rational problem-solving see section II of ch. 3; see also Maxwell (1984a, pp. 67-75 or 2007, pp. 80-88).

which become relevant when we are confronted by some especially recalcitrant problem – such as the problem of understanding the nature of the universe, or the problem of creating a civilized world – which can only be solved gradually and progressively, bit by bit, and not all at once.¹⁶⁹ Popper was too hostile to specialization to emphasize the need for rule 3; he did not appreciate that the evils of specialization can be counteracted by implementing rule 4.

It might seem that in moving from scientific method to critical and problem-solving rationality we lose the idea of learning from experience; but this is not so. Problem-solving rationality, as enshrined in the above four rules, is a method of learning from experience. Experience is what we acquire through trying out various possible solutions to the problem we wish to solve, and discovering that these possibilities more or less fail. Consider, for example, a problem of action, a technological or political problem, perhaps: in criticizing a proposed solution we may well appeal to the (adverse) outcome of attempting to put the solution into practice; that is, we appeal to experience. Experience, in this broad sense, is what we acquire through trying to do things, trying to solve problems: it is a generalization of the notion of experience as this arises in connection with science – observation and experimentation. Problem-solving rationality might also be called "problem-solving empiricism"; it is as much a generalization of scientific empiricism as it is of scientific rationality.

These four rules, though by no means sufficient for rationality,¹⁷⁰ are certainly necessary for it. No mode of inquiry can

¹⁶⁹. Can problem-solving rationality be regarded as a generalization of Popper's falsificationist conception of scientific method? In Popper (1959, 276-8) there is a discussion of the "'inductive' direction" of the growth of science: science proceeds by putting forward laws and theories of ever greater empirical content. This could be regarded as a case of putting rule 3 into practice. But rule 2 cannot be put into practice in science, in the context of "justification", since this would involve considering metaphysical theses as a part of scientific knowledge (as Popper himself in effect notes), which goes against Popper's demarcation criterion. This point comes up again in section 5 of the paper.

¹⁷⁰. These rules are not sufficient for rationality in part because of a lack of specific detail about how to improve aims and methods when aims are

hope to be rational which systematically violates any of these rules. In a moment we shall see that academic inquiry as it exists in the main at present, devoted to the pursuit of knowledge, systematically violates *three* of these four elementary, almost banal, entirely uncontroversial, rules of reason.

Two preliminary points now need to be made.

First, in order to create a more civilized, enlightened world, the problems that we need to solve are, fundamentally, problems of *living* rather than problems of *knowledge*. It is what we *do* (or refrain from doing) that matters, and not just what we *know*. Even where new knowledge or technology is needed, in connection with agriculture or medicine for example, it is always what this enables us to *do* that solves the problem of living.

Second, in order to make progress towards a sustainable, civilized world we need to learn how to resolve our conflicts in more cooperative ways than at present. A group acts cooperatively in so far as all members of the group share responsibility for what is done, and for deciding what is done, proposals for action, for resolution of problems and conflicts, being judged on their merits from the standpoint of the interests of the members of the group (or the group as a whole), there being no permanent leadership or delegation of power.¹⁷¹ Competition is not opposed to cooperation if it proceeds within a framework of cooperation, as it does ideally within science. There are of course degrees of cooperativeness, from its absence, all out violence, at one extreme, through settling of conflicts by means of threat, agreed procedures such as voting, via bargaining, to all out cooperativeness at the other extreme. If we are to develop a sustainable, civilized world we need to move progressively away from the violent end of this spectrum towards the cooperative end.

problematic, and in part because the list of rules is by no means complete (see Maxwell, 1984a, 69-75). A very important additional rule is:

5. In seeking to solve a problem, P*, search for an analogous, already solved problem, P; if such a problem is found, modify the solution, S, appropriately, taking the similarities and differences between P and P* into account, so that S becomes S*, and consider this as a candidate solution to P*.

¹⁷¹. As I am using the term, a conflict is only resolved "cooperat-ively" if it is resolved "justly".

Granted, then, that the task of academic inquiry is to put the four rules of problem-solving rationality into practice in such a way as to help humanity learn how to make progress towards a civilized, enlightened world, the primary intellectual tasks must be:

1. To articulate, and try to improve the articulation of, those social problems of living we need to solve in order to make progress towards a better world.

2. To propose and critically assess possible, and actual, increasingly cooperative social actions – these actions to be assessed for their capacity to resolve human problems and conflicts, thus enhancing the quality of human life.

These intellectually fundamental tasks are undertaken by social inquiry, at the heart of the academic enterprise. Social inquiry also has the task of promoting increasingly cooperatively rational tackling of problems of living in the social world – in such contexts as politics, commerce, international affairs, industry, agriculture, the media, the law, education.

Academic inquiry also needs, of course, to implement the third rule of rational problem solving; that is, it needs:

3. To break up the basic problems of living into preliminary, simpler, analogous, subordinate, specialized problems of knowledge and technology, in an attempt to work gradually towards solutions to the basic problems of living.

But, in order to ensure that specialized and basic problem solving keep in contact with one another, the fourth rule of rational problem solving also needs to be implemented: that is, academic inquiry needs:

4. To interconnect attempts to solve basic and specialized problems, so that basic problem-solving may guide, and be guided by, specialized problem-solving.

In diagram 1 I have tried to depict the kind of inquiry that would emerge as a result of putting the above four rules of rational problem solving into academic practice, as just indicated. I will give some further details below.

There are a number of points to note about this "rational problem solving" conception of academic inquiry. Social inquiry is not, primarily, social science; it has, rather, the intellectually basic task of engaging in, and promoting in the social world, increasingly

cooperatively rational tackling of conflicts and problems of living.¹⁷² Social inquiry, so conceived, is actually intellectually more fundamental than natural science (which seeks to solve subordinate problems of knowledge and understanding). Academic inquiry, in seeking to promote cooperatively rational problem solving in the social world, must engage in a two-way exchange of ideas, arguments, experiences and information with the social world. The thinking, the problem solving, that really matters, that is really fundamental, is the thinking that we engage in, individually, socially and institutionally, as we live; the whole of academic inquiry is, in a sense, a specialized part of this, created in accordance with rule 3, but also being required to implement rule 4 (so that social and academic problem solving may influence each other). Academic inquiry, on this model, is a kind of peoples' civil service, doing openly for the public what actual civil services are supposed to do, in secret, for governments. Academic inquiry needs just sufficient power to retain its independence, to resist pressures from government, industry, the media, religious authorities, and public opinion, but no more. Academia proposes to, argues with, learns from, attempts to teach, and criticizes all sectors of the social world, but does not instruct or dictate. It is an intellectual resource for the public, not an intellectual bully.

The basic intellectual aim of inquiry may be said to be, not knowledge, but wisdom – wisdom being understood to be the desire, the active endeavour and the capacity to realize¹⁷³ what is desirable and of value in life, for oneself and others. Wisdom includes knowledge, know-how and understanding but goes beyond them in also including the desire and active striving for what is of value, the ability to experience value, actually and potentially, in the circumstances of life, the capacity to help realize what is of value for oneself and others, the capacity to help solve those problems of living that need to be solved if what is of value is to be realized, the capacity to use and develop knowledge,

¹⁷². Social inquiry needs, of course, to tackle problems of knowledge of the social world subordinate to the tackling of social problems of living, in accordance with rule 3.

¹⁷³. "Realize" is intentionally ambiguous in that it here means *both* "to apprehend" and "to make real".

technology and understanding as needed for the realization of value. Wisdom, like knowledge, can be conceived of not only in personal terms but also in institutional or social terms. Thus, the basic aim of academic inquiry, according to the view being indicated here, is to help us develop wiser ways of living, wiser institutions, customs and social relations, a wiser world.

So far academic inquiry has been characterized as having the task of helping humanity learn how to tackle its problems of living more rationally; nothing has been said about learning from experience. But, as I indicated above, the four rules of reason that we are considering are also rules for learning from experience; this has a vital role to play in the conception of inquiry we are considering. What we learn as a result of attempting to put into practice some proposed solution to a problem of living is of course all important for learning how to build a better world. A vital task for academic inquiry (especially for history) is to monitor the successes and failures of our past attempts at solving problems of living. As far as possible we should try to ensure that our failed social experiments, our failed attempts at solving social problems, are performed only *in imagination*, and not *in practice* in the real world, so that we only suffer the consequences of failure in imagination, and not in reality. But however vivid, far-seeing and accurate our imagination may be, failure in practice will always happen, and we should seek to learn all we can from it for future actions. To this extent, the conception of inquiry we are considering can be regarded as a kind of empiricism. In two crucial respects, however, it differs from what is usually meant by empiricism. First, what is learned is how to do things, how to realize what is of value, how to live, and not, primarily, what we learn in the context of science: knowledge of fact. And secondly, as I have already remarked, "experience" means something like "what we acquire as a result attempting to do things, attempting to realize what is of value", and not, primarily, what it means in the context of science: observation and experiment. (This latter meaning is a specialized version of the former meaning.)

It is important to appreciate that the conception of academic inquiry that we are considering is designed to help us to see, to know and to understand, for their own sake, just as much as it is

designed to help us solve practical problems of living. It might seem that social inquiry, in articulating problems of living and proposing possible solutions, has only a severely practical purpose. But engaging in this intellectual activity of articulating personal and social problems of living is just what we need to do if we are to develop a good empathic or "personalistic" understanding of our fellow human beings (and of ourselves) – a kind of understanding that can do justice to our humanity, to what is of value, potentially and actually, in our lives. In order to understand another person *as a person* (as opposed to a biological or physical system) I need to be able, in imagination, to see, desire, fear, believe, experience and suffer what the other person sees, desires, etc. I need to be able, in imagination, to enter into the other person's world; that is, I need to be able to understand his problems of living as he understands them, and I need also, perhaps, to understand a more objective version of these problems. In giving intellectual priority to the tasks of articulating problems of living and exploring possible solutions, social inquiry thereby gives intellectual priority to the development of a kind of understanding that people can acquire of one another that is of great intrinsic value. In my view, indeed, personalistic understanding is essential to the development of our humanity, even to the development of consciousness. Our being able to understand each other in this way is also essential for cooperatively rational action.

And it is essential for science. It is only because scientists can enter imaginatively into each other's problems and research projects that objective scientific knowledge can develop. At least two rather different motives exist for trying to see the world as another sees it: one may seek to improve one's knowledge of the other person; or one may seek to improve one's knowledge of the world, it being possible that the other person has something to contribute to one's own knowledge. Scientific knowledge arises as a result of the latter use of personalistic understanding – scientific knowledge being, in part, the product of endless acts of personalistic understanding between scientists (with the personalistic element largely suppressed so that it becomes

invisible). It is hardly too much to say that almost all that is of value in human life is based on personalistic understanding.¹⁷⁴

The basic intellectual aim of the kind of inquiry we are considering is to devote reason to the discovery of what is of value in life. This immediately carries with it the consequence that the arts have a vital *rational* contribution to make to inquiry, as revelations of value, as imaginative explorations of possibilities, desirable or disastrous, or as vehicles for the criticism of fraudulent values through comedy, satire or tragedy. Literature and drama also have a rational role to play in enhancing our ability to understand others personalistically, as a result of identifying imaginatively with fictional characters – literature in this respect merging into biography, documentary and history. Literary criticism bridges the gap between literature and social inquiry, and is more concerned with the *content* of literature than the means by which it achieves its effects.

Another important consequence flows from the point that the basic aim of inquiry is to help us discover what is of value, namely that our feelings and desires have a vital rational role to play within the intellectual domain of inquiry. If we are to discover for ourselves what is of value, then we must attend to our feelings and desires. But not everything that feels good is good, and not everything that we desire is desirable. Rationality requires that feelings and desires take fact, knowledge and logic into account, just as it requires that priorities for scientific research take feelings and desires into account. In insisting on this kind of interplay between feelings and desires on the one hand, knowledge and understanding on the other, the conception of inquiry that we are considering resolves the conflict between rationalism and romanticism, and helps us to acquire what we need if we are to contribute to building civilization: mindful hearts and heartfelt minds.

¹⁷⁴. For a more detailed discussion of the nature, significance, and intellectually fundamental character of "personalistic" understanding, and its role in "wisdom-inquiry", see Maxwell (1984a, 172-89 and 264-75); (2001a, chs. 5-7).

This, then, in bare outline, is the kind of academic inquiry that would have emerged from the 18th century Enlightenment if the third great blunder of the Enlightenment had not been made.

But the blunder was made. Instead of the progress-achieving methods of science (generalized to become problem-solving rationality) being applied to *social life*, scientific method was applied to the task of developing *social science* alongside natural science. The outcome is what we have (by and large) today: a kind of inquiry that gives intellectual priority to the task of acquiring *knowledge*, this knowledge, once acquired, being subsequently and secondarily applied to help solve social problems. Rule 3 of problem solving rationality is put into practice to splendid effect: the outcome is the maze of specialized disciplines of the formal, natural, social and technological sciences that go to make up much of academic inquiry today. But rules 1, 2 and 4 are violated. Academic inquiry today, restricted primarily to solving problems of knowledge, is so irrational, in a wholesale and structural way, that *three* of the four most elementary rules of reason conceivable are violated. Rule 1 is violated because academia can articulate problems of *knowledge* but cannot, at a fundamental level, articulate problems of *living*. Rule 2 is violated because academia can propose and critically assess possible solutions to problems of *knowledge* – theories, observational and experimental results, factual claims of all kinds – but cannot propose and critically assess possible solutions to problems of living – proposals for action, policies, political programmes, political philosophies, philosophies of life. All these latter do not state matters of fact; they embody proposals as to what we should do, how we should live, what we should seek to change and create; they incorporate such things as values, human hopes and fears, policies, strategies for living: they do not constitute potential contributions to knowledge, and are thus excluded from a kind of inquiry devoted to the pursuit of knowledge. Once rules 1 and 2 are violated, rule 4 is necessarily violated as well.

This wholesale, structural irrationality of academic inquiry as it mostly exists today is no mere formal matter. It has far-flung, long-term damaging consequences. It means that knowledge and technological know-how are pursued dissociated intellectually

from a more fundamental concern to promote increasingly cooperatively rational tackling of conflicts and problems of living. As I have pointed out (in section 2), it is this that is at the root of most of our current global problems.

A kind of inquiry that pursues knowledge and technological know-how, and fails to give intellectual priority to the tasks of articulating our problems of living, and proposing and criticizing possible solutions (thus violating three of the four most elementary rules of reason conceivable), must inevitably tend to *create* the kind of global problems we face today, the outcome of possessing much recently acquired power to act without the power to act wisely. And the more successful such "knowledge-inquiry" is, so the greater the human suffering it is likely to lead to. Reason is far too important "for everything we cherish" for it to be tolerable that it should be systematically violated in this way.

Here, then, is a major failing of Popper's version the Enlightenment Programme. Instead of arguing for the need to reject "knowledge-inquiry" and replace it with the kind of academic inquiry indicated above, "wisdom-inquiry" as it may be called, Popper defends "knowledge-inquiry", and even defends pro-naturalist social science.

5 The New Enlightenment, Step (i): From Falsificationism to Aim-Oriented Empiricism

I come now to a rather more radical revision of Popper's version of the Enlightenment Programme. This begins with a revision of step (i) of the Programme. Popper's falsificationism is untenable, and needs to be replaced by a conception of scientific method which I have called *aim-oriented empiricism* (AOE). The reason for this revision can be summarized as follows. Science only considers (and only accepts) theories that are sufficiently simple, unified or explanatory, and this means that the methods of science make a persistent metaphysical assumption about the universe, to the effect that it has a simple, unified, explanatory dynamic structure. That such a persistent metaphysical assumption is made by science, as a part of (conjectural) scientific knowledge, contradicts, and refutes, falsificationism. An improved conception of scientific method is required.

In *The Logic of Scientific Discovery*, Popper claims that the more falsifiable a theory is, so the greater its degree of simplicity. (There is a second method for assessing degrees of simplicity, in terms of number of observation statements required to falsify the theories in question, but Popper stresses that if the two methods clash, it is the first that takes precedence.) It is easy to see that Popper's proposal fails. Given a reasonably simple scientific theory, T, one can readily increase the falsifiability of T by adding on an independently testable hypotheses, h_1 , to form the new theory, $T + h_1$. This new theory will be more falsifiable than T but, in general, will be drastically less simple. And one can make the situation even worse, by adding on as many independently testable hypotheses as one pleases, h_2, h_3 and so on, to form new theories $T + h_1 + h_2 + h_3 + \dots$, as highly empirically falsifiable and as drastically lacking in simplicity, as one pleases.¹⁷⁵ Thus simplicity cannot be equated with falsifiability.

¹⁷⁵. Given any accepted scientific theory, whether Newtonian theory, general relativity, quantum theory or the standard model, endlessly many rival theories can be concocted in each case that are even more falsifiable, better corroborated but, if anything, even more seriously *ad hoc* (i.e. lacking in simplicity) than the accepted theory, in the following way. Taking Newtonian theory (NT) as an example of an accepted theory, here are two examples of grossly *ad hoc* rival theories. NT*: "Everything occurs as NT asserts, until the first second of 2100, when an inverse cube law of gravitation will abruptly hold". NT**: "Everything occurs as NT asserts, except for systems consisting of gold spheres, each having a mass of 1,000 tons, interacting with each other gravitationally in outer space, in a vacuum, within a spherical region of 10 miles: for these systems, Newton's law of gravitation is repulsive, not attractive". It is easy to see that there are infinitely many such rivals to NT, all just as empirically successful (at the moment) as NT. The predictions of NT may be represented as points in a multi-dimensional space, each point corresponding to some specific kind of system (there being infinitely many points). NT has only been corroborated for a minute region of this space. In order to concoct a (grossly *ad hoc*) rival to NT, just as well corroborated as NT, all we need do is identify some region in this space that includes no prediction of NT that has been verified, and then modify the laws of NT arbitrarily, for just that identified region. Rival theories, of the above type, can easily be concocted that satisfy falsificationist requirements for being more acceptable than NT. NT, like most accepted physical theories, yields predictions that clash with observation or experiment, and thus are ostensibly falsified. We can always concoct new theories, in the way just indicated, doctored to yield the "correct" predictions. We can add on independently testable auxiliary postulates,

And there is a further, even more devastating point. Popper's methodological rules favour $T + h_1 + h_2 + h_3$ over T , especially if h_1 , h_2 and h_3 have been severely tested, and corroborated. But in scientific practice, $T + h_1 + h_2 + h_3$ would never even be considered, however highly corroborated it might be if considered, because of its extreme lack of simplicity or unity, its grossly *ad hoc* character. There is here a fundamental flaw in the central doctrine of *The Logic of Scientific Discovery*.

Later, in *Conjectures and Refutations* (1963), Popper put forward a new methodological principle which, when added to those of the earlier book, succeeds in excluding theories such as $T + h_1 + h_2 + h_3$ from scientific consideration. This principle states that a new theory, in order to be acceptable, "should proceed from some *simple, new, and powerful, unifying idea* about some connection or relation (such as gravitational attraction) between hitherto unconnected things (such as planets and apples) or facts (such as inertial and gravitational mass) or new "theoretical entities" (such as field and particles)" (Popper, 1963, 241). $T + h_1 + h_2 + h_3$ does not "proceed from some *simple, new and powerful, unifying idea*" and is to be rejected on that account, even if more highly corroborated than T .

But the adoption of this "requirement of simplicity" (as Popper calls it) as a basic methodological principle of science has the effect of permanently excluding from science all *ad hoc* theories (such as $T + h_1 + h_2 + h_3$) that fail to satisfy the principle, however empirically successful such theories might be if considered. This amounts to assuming permanently that the universe is such that no *ad hoc* theory, that fails to satisfy Popper's principle of simplicity, is true. It amounts to accepting, as a permanent item of scientific knowledge, the substantial metaphysical thesis that the universe is non-*ad hoc*, in the sense that no theory that fails to satisfy Popper's principle of simplicity is true. But this clashes with Popper's

thus ensuring that the new theory has greater empirical content than the old one. And no doubt this excess content will be corroborated. For further examples and discussion, see Maxwell (1998, 47-54).

criterion of demarcation: that no unfalsifiable, metaphysical thesis is to be accepted as a part of scientific knowledge.¹⁷⁶

It is, in fact, important that Popper's criterion of demarcation is rejected, and the metaphysical thesis of non-*ad hocness* is explicitly acknowledged to be a part of scientific knowledge. The thesis, in the form in which it is implicitly adopted at any given stage in the development of science, may well be false. Scientific progress may require that it be modified. The thesis needs to be made explicit, in other words, for good Popperian reasons, namely, so that it can be critically assessed, and perhaps improved. As long as Popper's demarcation criterion is upheld, the metaphysical thesis must remain implicit, and hence immune to criticism.

Popper's falsificationism can be modified, however, so that substantial metaphysical theses, implicit in methods that exclude *ad hoc* theories, are made explicit within science, and are acknowledged to be basic items of (conjectural) scientific knowledge, thus becoming open to critical scrutiny and revision. The outcome is a more rational, a more intellectually rigorous kind of science, just because substantial, influential and problematic metaphysical theses, implicit in the methods of science, become explicitly criticizable and improvable.

The moment we acknowledged that there is a persistent metaphysical thesis implicit in the methods of science, two new problems leap to our attention. What, precisely, does this metaphysical thesis assert? And on what grounds is it to be

¹⁷⁶. But does implementing Popper's methodological "principle of simplicity" really commit science to the metaphysical thesis that the universe is simple? Suppose, instead of adopting Popper's principle, science adopted the principle: in order to be acceptable, a new physical theory must postulate that the universe is made up of atoms. This methodological principle is upheld in such a way that even though theories are available which postulate fields rather than atoms, and which are much more empirically successful than any atomic theory, nevertheless these rival field theories are all excluded from science. Would it not be clear that science, in adopting and implementing the methodological principle of atomicity in this way, is making the assumption that the universe is made up of atoms, whether this is acknowledged or not? How can this be denied? Just the same holds if science adopts and implements Popper's methodological principle of simplicity.

(conjecturally) accepted as a part of scientific knowledge? AOE is put forward as the solution to these two problems.

As far as the first of the above two problems is concerned, a wide range of metaphysical theses are available. At one extreme, we might adopt a metaphysical thesis that excludes only utterly silly theories; at the other extreme, we might adopt the thesis that the universe is physically comprehensible in the sense that it has a unified dynamic structure, some yet-to-be-discovered unified physical "theory of everything" being true – a thesis that I shall call "physicalism". We might even adopt some specific version of physicalism, which asserts that the underlying physical unity is of a specific type: it is made up of a unified field perhaps, a quantum field, or empty topologically complex curved space-time, or a quantum string field. Other things being equal, the more specific the thesis (and thus the more it excludes) so the more likely it is to be false, whereas the more unspecific it is so the more likely it is to be true.

As far as the second of the above two problems is concerned, there are three considerations that we can appeal to, wholly Popperian in spirit if not in the letter of Popperian doctrine.

(1) If some metaphysical thesis, M , is implicit in some scientific methodological practice, then science is more rigorous if M is made explicit, since this facilitates criticism of it, the consideration of alternatives.

(2) A metaphysical thesis may be such that its truth is a necessary condition for it to be possible for us to acquire knowledge: if so, accepting the thesis can only help, and cannot undermine, the pursuit of knowledge of truth.

(3) Given two rival metaphysical theses, M_1 and M_2 , it may be the case that M_1 supports an empirical scientific research programme that has apparently met with far greater empirical success than any rival empirical research programme based on M_2 : in this case we may favour M_1 over M_2 , at least until M_2 , or some third thesis, M_3 , shows signs of supporting an even more empirically progressive research programme.

Two difficulties arise, however, when one attempts to use (2) and (3) to select the best available metaphysical thesis from the infinitely many options available. As far as (2) is concerned, any

thesis sufficiently substantial to exclude empirically successful crackpot theories from science is such that acquisition of knowledge might still be possible even if the thesis is false. On the other hand, any thesis such that its truth is necessary for knowledge to be acquired is much too insubstantial to exclude crackpot theories. As far as (3) is concerned, given any metaphysical thesis, M, that supports a non-crackpot empirically progressive scientific research programme, we can mimic this with a crackpot M* that supports a crackpot empirically progressive research programme, with a series of crackpot theories, T1*, T2*, ..., these theories becoming progressively more and more empirically successful, and closer and closer to exemplifying M*.

These two difficulties can be overcome, however, if physics is construed as adopting a hierarchy of metaphysical conjectures concerning the comprehensibility and knowability of the universe, these conjectures becoming more and more insubstantial as one ascends the hierarchy, more and more likely to be true: see diagram 2. As I have formulated this idea, in Maxwell (1998), the top conjecture in the hierarchy (at level 10) asserts, roughly, that the universe is such that some (conjectural) knowledge of our local circumstances can be acquired. This, and the next conjecture down are, I argue, to be accepted as permanent items of scientific knowledge, in accordance with (2), on the grounds, that is, that such acceptance can only help, and cannot hinder, the search for factual knowledge whatever the universe may be like. At level 4 the conjecture to be adopted is, I argue, physicalism. At level 5 there is the less precise conjecture that the universe is comprehensible in some way or other; and as one goes up the hierarchy to levels 9 and 10, the conjectures become progressively less and less substantial and precise. At level 3 there is the best currently available more or less specific version of physicalism, which I call the current "metaphysical blueprint". Examples from the history of physics are: the universe consists of (a) corpuscles which interact by contact (b) point-atoms which interact by means of forces (c) a unified classical field (d) a unified quantum field (e) empty, curved, topologically complex space-time (f) a unified quantum string field. At level 2 are currently accepted fundamental physical theories, and at level 1 there are empirical data. Two

considerations govern acceptance of metaphysical conjectures from level 3 to level 8. Any such conjecture must, as far as possible (A) exemplify, be a precise version of, and imply, the next conjecture up in the hierarchy, (B) be more empirically fruitful than any rival conjecture, in that it is a part of an empirical research programme that seems to be more empirically progressive than any rival research programme, in accordance with (3) above. Two considerations also govern acceptance of testable fundamental dynamical physical theories. Such a theory must be such that (i) it, together with all other accepted fundamental physical theories, exemplifies, or is a special case of, the best available metaphysical blueprint (at level 3) to a sufficiently good extent, (ii) it is sufficiently successful empirically (where empirical success is to be understood, roughly, in a Popperian sense).

This hierarchical view of AOE overcomes the two difficulties, indicated above, roughly as follows. Only the top two theses are accepted as a result of an appeal to (2); theses at levels 3 to 8 are accepted as a result of (a) an appeal to (3), and (b) compatibility with the top two theses at levels 10 and 9; this suffices to exclude aberrant rivals at levels 3 to 8 (which might be construed to support aberrant, empirically progressive research programmes). For further details of how AOE overcomes the two difficulties indicated above, and for further details of the view itself, see Maxwell (1998, chapter 5, and elsewhere).

A basic idea of AOE is to channel or direct criticism so that it is as fruitful as possible, from the standpoint of aiding progress in knowledge. The function of criticism within science is to promote scientific progress. When criticism demonstrably cannot help promote scientific progress, it becomes irrational (the idea behind (2) above). In an attempt to make criticism as fruitful as possible, we need to try to direct it at targets which are the most fruitful, the most productive, to criticize (from the standpoint of the growth of knowledge). This is the basic idea behind the hierarchy of AOE. Conjectures at all levels remain open to criticism. But, as we ascend the hierarchy, conjectures are less and less likely to be false; it is less and less likely that criticism, here, will help promote

scientific knowledge.¹⁷⁷ The best currently available level 3 conjecture is almost bound to be false: the history of physics reveals, at this level, as I have indicated above, that a number of different conjectures have been adopted and rejected in turn. Here, criticism, the activity of developing alternatives (compatible with physicalism) is likely to be immensely fruitful for progress in theoretical physics. Indeed, in Maxwell (1998, 78-89, 159-163 and especially 217-223), I argue that this provides physics with a rational, though fallible and non-mechanical method for the discovery of new fundamental physical theories, a method invented and exploited by Einstein in discovering special and general relativity (Maxwell, 1993, 275-305), something which Popper has argued is not possible: see Popper (1959, 31-2). Criticizing physicalism, at level 4, may also be fruitful for physics, but (the conjecture of AOE is) that this is not as likely to be as fruitful as criticism at level 3. (Elsewhere I have suggested alternatives to physicalism.) And, as we ascend the hierarchy (so AOE conjectures), criticism becomes progressively less and less likely to be fruitful. Against that, it must be admitted that the higher in the hierarchy we need to modify our ideas, so the more dramatic the intellectual revolution that this would bring about. If physicalism is rejected altogether, and some quite different version of the level 5 conjecture of comprehensibility is adopted instead, the whole character of natural science would change dramatically; physics, as we know it, might even cease to exist.

The biggest change, in moving from falsificationism to AOE, has to do with the role of metaphysics in science, and the scope of scientific knowledge. According to falsificationism, untestable metaphysical theses may influence scientific research in the context of discovery, and may even lead to metaphysical research programmes; they cannot, however, be a part of scientific knowledge itself. But according to AOE, the metaphysical theses at levels 3 to 10 are all a part of current (conjectural) scientific knowledge. In particular, physicalism is. According to AOE, it is a

¹⁷⁷. It is not enough just to be critical: we need to be *critically* critical. Criticism needs to be directed at those points in a possible solution to a problem where it is most likely to be fruitful.

part of current (conjectural) scientific knowledge that the universe is physically comprehensible – certainly not the case granted falsificationism.

Another important change has to do with the relationship between science and the philosophy of science. Falsificationism places the study of scientific method, the philosophy of science, outside science itself, in accordance with Popper's demarcation principle. AOE, by contrast, makes scientific method and the philosophy of science an integral part of science itself. The activity of tackling problems inherent in the aims of science, at a variety of levels, and of developing new possible aims and methods, new possible more specific or less specific philosophies of science (views about what the aims and methods of science ought to be) is, according to AOE, a vital research activity of science itself. But this is also philosophy of science, being carried on within the framework of AOE.¹⁷⁸

AOE differs in many other important ways from Popper's falsificationism, (see Maxwell, 1998). Nevertheless the impulse, the intellectual aspirations and values, behind the hierarchical view of AOE are, as I have tried to indicate, thoroughly Popperian in character and spirit. The whole idea is to turn implicit assumptions into explicit conjectures in such a way that criticism may be directed at what most needs to be criticized from the standpoint of aiding progress in knowledge, so that conjectures may be developed and adopted that are the most fruitful in promoting scientific progress, at the same time no substantial conjecture, implicit or explicit, being held immune from critical scrutiny. AOE provides a framework of relatively unspecific, unproblematic aims

¹⁷⁸. In holding that metaphysical theses and philosophies of science are an integral part of science itself, AOE implies that Popper's principle of demarcation (Popper, 1963, chapter 11) is to be rejected. Popper's demarcation proposal, apart from being untenable, is in any case too simplistic, in that it reduces to one a number of distinct demarcation issues. Popper rolls into one the distinct tasks of demarcating (a) good from bad science, (b) science from non-science, (c) science from pseudo-science, (d) rational from irrational inquiry, (e) knowledge from mere speculation, (f) knowledge from dogma (or superstition, or prejudice, or popular belief), (g) the empirical from the metaphysical, and (h) factual truth from non-factual (analytic) truth. (a) to (d) involve demarcating between disciplines, whereas (e) to (h) involve demarcating between propositions.

and methods within which much more specific, problematic aims and methods may be improved as our knowledge improves. The hope is that as we increase our knowledge about the world we improve the cosmological assumptions implicit in our methods, and thus in turn improve our methods. As a result of improving our knowledge we improve our knowledge about how to improve knowledge. There is a kind of positive feedback between improving knowledge, and improving aims and methods. Science adapts its own nature to what it learns about the nature of the universe, thus increasing its capacity to make progress in knowledge about the world – the methodological key to the astonishing, accelerating, explosive growth of scientific knowledge.¹⁷⁹

6 The New Enlightenment, Step (ii): From Critical to Aim-Oriented Rationalism

Falsificationism is defective because it fails to identify the problematic aim of science properly, and thus fails to specify the need for science to improve its aims and methods as it proceeds. Critical rationalism is defective in an analogous way. It does not make improving aims and methods, when aims are problematic, an essential aspect of rationality.

If, however, we take AOE as our starting point, and generalize that, the outcome is different. It is not just in science that aims are problematic; this is the case in life too, either because different aims conflict, or because what we believe to be desirable and realizable lacks one or other of these features, or both. Above all, the aim of creating global civilization is inherently and profoundly problematic. Furthermore, it is not just science that problematic aims are "repressed" (see Maxwell, 2002a); many other institutional and traditional endeavours repress problematic aims and acknowledge ostensibly unproblematic, token aims instead.

¹⁷⁹. For a detailed exposition of AOE, an account of how it solves problems of simplicity, unity, induction and verisimilitude, and criticism of rival views, see Maxwell (1998). AOE has emerged from criticism of falsificationism over many years: see Maxwell (1972a, 1974, 1976a, 1977a, 1979, 1984a, ch. 9, 1993). For recent brief accounts of the case for AOE see Maxwell (1997, 1999a, 2000b, 2002b).

Quite generally, then, and not just in science, whenever we pursue a problematic aim we need first, to acknowledge the aim; then we need to represent it as a hierarchy of aims, from the specific and problematic at the bottom of the hierarchy, to the general and unproblematic at the top. In this way we provide ourselves with a framework within which we may improve more or less specific and problematic aims and methods as we proceed, learning from success and failure in practice what it is that is both of most value and realizable. Such an "aim-oriented" conception of rationality is the proper generalization of the aim-oriented, progress-achieving methods of science.¹⁸⁰

Any conception of rationality which systematically leads us astray must be defective. But any conception of rationality, such as Popper's critical rationalism, which does not include explicit instructions for the *improvement* of aims, must systematically lead us astray. It will do so whenever we fail to choose that aim that is in our best interests or, more seriously, whenever we misrepresent our aim – as we are likely to do whenever aims are problematic. In these circumstances, the more "rationally" we pursue the aim we acknowledge, the worse off we will be. Systematically, such conceptions of rationality, which do not include provisions for improving problematic aims, are a hinderance rather than a help; they are, in short, defective.¹⁸¹

AOE and its generalization, aim-oriented rationality (AOR), incorporate all the good points of Popper's falsificationist conception of science and its generalization, critical rationalism, indicated above, but also improve on Popper's notions, in being designed to help science and other worthwhile endeavours progressively improve problematic aims and methods.

7 The New Enlightenment, Step (iii): From Knowledge to Wisdom

I come now to step (iii) of the new Enlightenment Programme. The task, here, is to help humanity gradually get more AOR into

¹⁸⁰. See Maxwell (1976a, 1984a, ch. 5, and 2001a, ch. 9).

¹⁸¹. Science specifically, and academic inquiry more generally, misrepresent basic aims in just this way: see Maxwell (1984a, 2002a).

diverse aspects of social and institutional life – personal, political, economic, educational, international – so that humanity may gradually learn how to make progress towards an enlightened world. Social inquiry, in taking up this task, needs to be pursued as social *methodology* or social *philosophy*. What the philosophy of science is to science, as conceived by AOE, so sociology is to the social world: it has the task of helping diverse valuable human endeavours and institutions gradually improve aims and methods so that the world may make social progress towards global enlightenment. (The sociology of science, as a special case, is one and the same thing as the philosophy of science.) And a basic task of academic inquiry, more generally, becomes to help humanity solve its problems of living in increasingly rational, cooperative, enlightened ways, thus helping humanity become more civilized. The basic aim of academic inquiry becomes, I have already said, to promote the growth of *wisdom*. Those parts of academic inquiry devoted to improving knowledge, understanding and technological know-how contribute to the growth of wisdom. The New Enlightenment Programme thus has dramatic and far reaching implications for academic inquiry, for almost every branch and aspect of science and the humanities, for its overall character and structure, its overall aims and methods, and its relationship to the rest of the social world. I have spelled out in some detail what these implications are in a number of publications.¹⁸²

As I have already remarked, the aim of achieving global civilization is inherently problematic.¹⁸³ This means, according to

¹⁸². See ch. 2 and Maxwell (1976a, 1977a, 1984b, 1987, 1991, 1992a, 1992b, 1997b, 1998, 2000a, 2000b, 2001a, 2001b, 2002a, 2002b, and especially 1984a or 2007)

¹⁸³. Fundamentally, this is due to the profound difficulty of discovering what is *achievable* (by increasingly civilized means), and *of value*. People hold conflicting views about what is achievable and of value, all too often in a highly dogmatic way, ignoring the profoundly problematic character of the whole idea of civilization. Many well-known views that have been proposed as to what constitutes Utopia, an ideally civilized society, have been unrealizable, horrifically undesirable, or both, attempts to realize such ideals, when taken up in practice, leading to various kinds of hell on earth (as in Hitler's Germany, Stalin's Russia, or Mao's China). Furthermore, it is not just that people have conflicting interests, values and ideals; even our very best ideas as to what constitutes

AOR, that we need to represent the aim at a number of levels, from the specific and highly problematic to the unspecific and unproblematic. Thus, at a fairly specific level, we might, for example, specify civilization to be a state of affairs in which there is an end to war, dictatorships, population growth, extreme inequalities of wealth, and the establishment of democratic, liberal world government and a sustainable world industry and agriculture. At a rather more general level we might specify civilization to be a state of affairs in which everyone shares equally in enjoying, sustaining and creating what is of value in life *in so far as this is possible*. Diagram 3 depicts a cartoon version of what is required, arrived at by generalizing and then reinterpreting diagram 2.

As a result of building into our institutions and social life such a hierarchical structure of aims and associated methods, we create a framework within which it becomes possible for us progressively to improve our real-life aims and methods in increasingly cooperative ways as we live. Diverse philosophies of life – diverse religious, political, economic and moral views – may be cooperatively developed, assessed and tested against the experience of personal and social life. It becomes possible progressively to improve diverse *philosophies of life* (diverse views about what is of value in life and how it is to be realized)

civilization embody (and need to embody) conflicting ideals. Thus freedom and equality, even though inter-related, may nevertheless clash. It would be an odd notion of individual freedom which held that freedom was for some, and not for others; and yet if equality is pursued too singlemindedly this will undermine individual freedom, and will even undermine equality, in that a privileged class will be required to enforce equality on the rest, as in the Soviet Union. A basic aim of legislation for civilization, we may well hold, ought to be increase freedom by restricting it: this brings out the inherently problematic character of the aim of achieving civilization. One thinker who has stressed the inherently contradictory character of the idea of civilization is Isaiah Berlin: see, for example, Berlin (1980, 74-9. Berlin thought the problem could not be solved, but this was because he was ignorant of aim-oriented rationality. In depicting ideals of civilization at a hierarchy of levels, aim-oriented rationality provides the means for progressively improving resolutions to inherently conflicting ideals, such as freedom and equality.

much as *theories* are progressively and cooperatively improved in science.

AOR is especially relevant when it comes to resolving conflicts cooperatively. If two groups have partly conflicting aims but wish to discover the best resolution of the conflict, AOR helps in requiring of those involved that they represent aims at a level of sufficient imprecision for agreement to be possible, thus creating an agreed framework within which disagreements may be explored and resolved. AOR cannot, of itself, combat non-cooperativeness, or induce a desire for cooperativeness; it can however facilitate the cooperative resolution of conflicts if the desire for this exists. In facilitating the cooperative resolution of conflicts in this way, AOR can, in the long term, encourage the desire for cooperation to grow (if only because it encourages belief in the possibility of cooperation).

8 Objections

I now consider, briefly, some objections that may be raised against my claim that *The New Enlightenment* improves on the Popperian version of the Enlightenment Programme.

It may be objected that the Traditional Enlightenment does not dominate current academic inquiry to the extent that I have assumed. But grounds for holding that it does are given in chapter six of my *From Knowledge to Wisdom*. There I looked at the following: (1) books about the modern university; (2) the philosophy and sociology of science; (3) statements of leading scientists; (4) Physics Abstracts; (5) Chemistry, Biology, Geo and Psychology Abstracts; (6) journal titles and contents; (7) books on economics, sociology and psychology; (8) philosophy. In 1984, the year *From Knowledge to Wisdom* was published, there can be no doubt whatsoever that the Traditional Enlightenment (or "the philosophy of knowledge" as I called it in the book) dominated academic inquiry.

Have things changed since then? The revolution advocated by *From Knowledge to Wisdom*, and argued for here, has not occurred. There is still, amongst the vast majority of academics today, no awareness at all that a more intellectually rigorous and humanly valuable kind of inquiry than that which we have at

present, exists as an option. In particular, social inquiry continues to be taught and pursued as social *science*, and not as social *methodology*. Recently I undertook an examination, at random, of thirty-four introductory books on sociology, published between 1985 and 1997. Sociology, typically, is defined as "the scientific study of human society and social interactions" (Tischler, 1996, 4), as "the *systematic, sceptical study of human society*" (Macionis and Plummer, 1997, 4), or as having as its basic aim "to understand human societies and the forces that have made them what they are" (Lenski et al., 1995, 5). Some books take issue with the idea that sociology is the *scientific* study of society, or protest at the male dominated nature of sociology (for example, Abbott and Wallace, 1990, 3 and 1). Nowhere did I find a hint of the idea that a primary task of sociology, or of social inquiry more generally, might be to help build into the fabric of social life progress-achieving methods, generalized from those of science, designed to help humanity resolve its conflicts and problems of living in more cooperatively rational ways than at present.

The tackling of problems of living rather than problems of knowledge does of course go on within the academic enterprise as it is at present constituted, within such disciplines as economics, development studies, policy studies, peace studies, medicine, agriculture, engineering, and elsewhere. But this does not tell against the point that the primary task of academic inquiry at present is, first, to acquire knowledge and technological know-how, and then, second, to apply it to help solve problems of living. It does not, in other words, tell against the point that it is the Traditional Enlightenment that is the dominant influence on the nature, the aims and methods, the whole character and structure of academic inquiry.

It may be objected that it is all to the good that the academic enterprise today does give priority to the pursuit of knowledge over the task of promoting wisdom and civilization. Before problems of living can be tackled rationally, knowledge must first be acquired.¹⁸⁴

¹⁸⁴. This is the objection that most academics will wish to raise against the conception of inquiry implied by the "Improved Popperian Enlightenment" and

I have six replies to this objection.

First, even if the objection were valid, it would still be vital for a kind of inquiry designed to help us build a better world to include rational exploration of problems of living, and to ensure that this guides priorities of scientific research (and is guided by the results of such research).

Second, the validity of the objection becomes dubious when we take into account the considerable success people met with in solving problems of living in a state of extreme ignorance, before the advent of science. We still today often arrive at solutions to problems of living in ignorance of relevant facts.

Third, the objection is not valid. In order to articulate problems of living and explore imaginatively and critically possible solutions (in accordance with Popper's conception of rationality) we need to be able to act in the world, imagine possible actions and share our imaginings with others: in so far as some common sense knowledge is implicit in all this, such knowledge is required to tackle rationally and successfully problems of living. But this does not mean that we must give intellectual priority to acquiring new relevant knowledge before we can be in a position to tackle rationally our problems of living.

Fourth, simply in order to have some idea of what kind of knowledge or know-how it is *relevant* for us to try to acquire, we must *first* have some provisional ideas as to what our problem of living is and what we might do to solve it. Articulating our problem of living and proposing and critically assessing possible solutions needs to be intellectually prior to acquiring relevant knowledge simply for this reason: we cannot know what new knowledge it is *relevant* for us to acquire until we have at least a preliminary idea as to what our problem of living is, and what we propose to do about it. A slight change in the way we construe our problem may lead to a drastic change in the kind of knowledge it is relevant to acquire: changing the way we construe problems of

the "New Enlightenment". It will be made by all those who hold that academic inquiry quite properly seeks to make a contribution to human welfare by, first, acquiring knowledge and then, secondarily, applying it to help solve human problems.

health, to include *prevention* of disease (and not just curing of disease) leads to a dramatic change in the kind of knowledge we need to acquire (importance of exercise, diet etc.). Including the importance of avoiding *pollution* in the problem of creating wealth by means of industrial development leads to the need to develop entirely new kinds of knowledge.

Fifth, relevant knowledge is often hard to acquire; it would be a disaster if we suspended life until it had been acquired. Knowledge of how our brains work is presumably highly relevant to all that we do but clearly, suspending rational tackling of problems of living until this relevant knowledge has been acquired would not be a sensible step to take. It would, in any case, make it impossible for us to acquire the relevant knowledge (since this requires scientists to act in doing research). Scientific research is itself a kind of action carried on in a state of relative ignorance.

Sixth, the capacity to act, to live, more or less successfully in the world, is more fundamental than (propositional) knowledge. Put in Rylean terms, 'knowing how' is more fundamental than 'knowing that'.¹⁸⁵ All our knowledge is but a development of our capacity to act. Dissociated from life, from action, knowledge stored in libraries is just paper and ink, devoid of meaning. In this sense, problems of living are more fundamental than problems of knowledge (which are but an aspect of problems of living); giving intellectual priority to problems of living quite properly reflects this point.¹⁸⁶ The point made above in section 4 deserves to be re-emphasized: a kind of inquiry that gives priority to tackling problems of knowledge over problems of living violates the most elementary requirements of rationality conceivable. If the basic task is to help humanity create a better world, then the problems that need to be solved are, primarily, problems of living, problems of action, not problems of knowledge. This means that to comply, merely, with Popper's conception of critical rationalism (or problem-solving rationality) discussed above, the basic intellectual

¹⁸⁵. Ryle (1949, ch. II).

¹⁸⁶. For a development of this point, see Maxwell (1984, 174-181). In some respects it accords with Popper's views on the biological and evolutionary origins of human thought: see, for example, Popper (1972, ch 7).

tasks need to be (1) to articulate problems of living, and (2) to propose and critically assess possible solutions, possible more or less cooperative human *actions*. (1) and (2) are excluded, or marginalized, by a kind of inquiry that gives priority to the task of solving problems of knowledge. And the result will be a kind of inquiry that fails to create a reservoir of imaginative and critically examined ideas for the resolution of problems of living, and instead develops knowledge often unrelated to, or even harmful to, our most basic human needs.

It may be objected that in employing AOR in an attempt to help create a more civilized world, in the way indicated above, the New Enlightenment falls foul of Popper's strictures against Utopian social engineering.¹⁸⁷ I have three replies to this objection. First, to the extent that piecemeal social engineering, of the kind advocated by Popper, is indeed the rational way to make progress towards a more civilized world, this will be advocated by the New Enlightenment. Second, when we take into account the unprecedented *global* nature of many of our most serious problems, indicated at the beginning of this essay (the outcome of solving the first great problem of learning but failing to solve the second), we may well doubt that piecemeal social engineering is sufficient. Third, Popper's distinction between piecemeal and Utopian social engineering is altogether too crude: it overlooks entirely what has been advocated here, aim-oriented rationalistic social engineering, with its emphasis on developing increasingly cooperatively rational resolutions of human conflicts and problems in full recognition of the inherently problematic nature of the aim of achieving greater civilization.¹⁸⁸

All those to any degree influenced by Romanticism and what Isaiah has called the counter-Enlightenment will object strongly to the idea that we should learn from scientific progress how to achieve social progress towards civilization; they will object strongly to the idea of allowing conceptions of rationality, stemming from science, to dominate in this way, and will object even more strongly to the idea, inherent in the New Enlightenment,

¹⁸⁷. Popper (1966a, vol. 1, ch. 9); (1961, 64-92).

¹⁸⁸. For further discussion see Maxwell (1984a, 189-98).

that we need to create a more aim-oriented rationalistic social world.

Directed at the Traditional Enlightenment, objections of this kind have some validity; but directed at the New Enlightenment, they have none. As I have emphasized elsewhere, AOR amounts to a synthesis of traditional rationalist and romantic ideals, and not to the triumph of the first over the second. In giving priority to the realization of what is of value in life, and in emphasizing that rationality demands that we seek to improve aims as we proceed, the New Enlightenment requires that rationality integrates traditional Rationalist and Romantic values and ideals of integrity. Imagination, emotion, desire, art, empathic understanding of people and culture, the imaginative exploration of aims and ideals, which tend to be repudiated as irrational by traditional Rationalism, but which are prized by Romanticism, are all essential ingredients of aim-oriented rationality. Far from crushing freedom, spontaneity, creativity and diversity, AOR is essential for the desirable flourishing of these things in life.¹⁸⁹

Finally, it may object that science is too different from political life for there to be anything worthwhile to be learnt from scientific success about how to achieve social progress towards civilization.¹⁹⁰ (a) In science there is a decisive procedure for eliminating ideas, namely, empirical refutation: nothing comparable obtains, or can obtain, in the political domain. (b) In science experiments or trials may be carried out relatively painlessly (except, perhaps, when new drugs are being given in live trials); in life, social experiments, in that they involve people, may cause much pain if they go wrong, and may be difficult to stop once started. (c) Scientific progress requires a number of highly intelligent and motivated people to pursue science on the behalf of the rest of us, funded by government and industry; social progress requires almost *everyone* to take part, including the stupid, the criminal, the mad or otherwise handicapped, the ill, the highly unmotivated; and in general there is no payment. (d)

¹⁸⁹. See Maxwell (1984a, 63-4, 85-91 and 117-118), for further discussion of this issue. See also Maxwell (1976a, especially chs. 1 and 8-10).

¹⁹⁰. N. Rescher, personal communication; Durant (1997).

Scientists, at a certain level, have an agreed, common objective: to improve knowledge. In life, people often have quite different or conflicting goals, and there is no general agreement as to what civilization ought to mean, or even whether it is desirable to pursue civilization in *any* sense. (e) Science is about fact, politics about value, the quality of life. This difference ensures that science has nothing to teach political action (for civilization). (f) Science is male-dominated, fiercely competitive, and at times terrifyingly impersonal (Harding, 1986); this means it is quite unfit to provide any kind of guide for life.

Here, briefly, are my replies. (a) Some proposals for action can be shown to be unacceptable quite decisively as a result of experience acquired through attempting to put the proposal into action. Where this is not possible, it may still be possible to assess the merits of the proposal to some extent by means of experience. If assessing proposals for action by means of experience is much more indecisive than assessing scientific theories by means of experiment, then we need, all the more, to devote our care and attention to the former case. (b) Precisely because experimentation in life is so much more difficult than in science, it is vital that in life we endeavour to learn as much as possible from (i) experiments that we perform in our imagination, and (ii) experiments that occur as a result of what actually happens. (c) Because humanity does not have the aptitude or desire for wisdom that scientists have for knowledge, it is unreasonable to suppose that progress towards global wisdom could be as explosively rapid as progress in science. Nevertheless progress in wisdom might go better than it does at present.

(d) Cooperative rationality is only feasible when there is the common desire of those involved to resolve conflicts in a cooperatively rational way. (e) Aim-oriented rationality can help us improve our decisions about what is desirable or of value, even if it cannot reach decisions for us. (f) In taking science as a guide for life, it is the progress-achieving methodology of science to which we need to attend. It is this that we need to generalize in such a way that it becomes fruitfully applicable, potentially, to all that we do. That modern science is male-dominated, fiercely competitive, and at times terrifyingly impersonal should not deter

us from seeing what can be learned from the progress-achieving methods of science – unless, perhaps, it should turn out that being male-dominated, fiercely competitive and impersonal is essential to scientific method and progress. (But this, I submit, is not the case.)

9 Implications for Academic Inquiry

Popper's version of the Enlightenment Programme, as enshrined in his first four books, has major implications, still unacknowledged by many, for a wide range of human endeavours, such as politics, education, the arts, philosophy, the humanities. But when developed further, in ways indicated above, the Popperian Enlightenment has even more fruitful, dramatic widespread implications. It is hardly too much to say, in my view, that the upshot of the argument is that we require a social and cultural revolution as substantial and dramatic, perhaps, as that of the Reformation, the Scientific Revolution, or the 18th century Enlightenment itself. This revolution involves changing the Traditional Enlightenment and the Romantic opposition so that these become unified in the New Enlightenment; it involves appropriately modifying all those activities and institutions affected by the Traditional Enlightenment and the Romantic opposition so that they come to embody the New Enlightenment: science, art, politics, education, medicine, philosophy, law, industry, agriculture, education.

In particular it involves changing academic inquiry so that, instead of being shaped by the Traditional Enlightenment (modified somewhat by Popper) and the Romantic opposition, as at present, it comes to be shaped by the New Enlightenment. At the end of the next chapter, I give a list of fifteen structural changes that need to be made to academic inquiry if it is to come to embody the New Enlightenment. The upshot would be universities rationally devoted to helping us realize what is of value in life, rationally devoted to helping us make progress towards as enlightened a world as possible.

Chapter Eight

Can The World Learn Wisdom?

(First published in *Solidarity, Sustainability, and Non-Violence*, vol. 3, no. 4, April 2007; www.pelicanweb.org/solisustv03n04maxwell.html)

The crisis of our times is that we have science without wisdom. This is the crisis behind all the others. Population growth, the terrifyingly lethal character of modern war and terrorism, immense differences of wealth across the globe, annihilation of indigenous people, cultures and languages, impending depletion of natural resources, destruction of tropical rain forests and other natural habitats, rapid mass extinction of species, pollution of sea, earth and air, thinning of the ozone layer, above all global warming - even the aids epidemic: all these relatively recent crises have been made possible by modern science and technology. Indeed, in a perfectly reasonable sense of "cause", they have been *caused* by modern science and technology.

It may be objected that it is not *science* that is the cause of these global problems but rather the things that we *do*, made possible by science and technology. This is obviously correct. But it is also correct to say that scientific and technological progress *is* the cause. The meaning of "cause" is ambiguous. By "the cause" of event E we may mean something like "the most obvious observable events preceding E that figure in the common sense explanation for the occurrence of E". In this sense, human actions (made possible by science) are the cause of such things as people being killed in war, destruction of tropical rain forests. On the other hand, by the "cause" of E we may mean "that prior change in the environment of E which led to the occurrence of E, and without which E would not have occurred". If we put the 20th century into the context of human history, then it is entirely correct to say that, in this sense, scientific-and-technological progress is the cause of our distinctive current global disasters: what has changed, what is

new, is scientific knowledge, not human nature. (Give a group of chimpanzees rifles and teach them how to use them and in one sense, of course, the cause of the subsequent demise of the group would be the actions of the chimpanzees. But in another obvious sense, the cause would be the sudden availability and use of rifles – the new, lethal technology.) Yet again, from the standpoint of theoretical physics, "the cause" of E might be interpreted to mean something like "the physical state of affairs prior to E, throughout a sufficiently large spatial region surrounding the place where E occurs". In this third sense, the sun continuing to shine is as much a part of the cause of war and pollution as human action or human science and technology.

In short, if by the cause of an event we mean that prior change which led to that event occurring, then it is the advent of modern science and technology that has caused all our current global crises. It is not that people became greedier or more wicked in the 19th and 20th centuries; nor is it that the new economic system of capitalism is responsible, as some historians and economists would have us believe. The crucial factor is the creation and immense success of modern science and technology. This has led to modern medicine and hygiene, to population growth, to modern agriculture and industry, to world wide travel (which spreads diseases such as aids), to global warming, and to the destructive might of the technology of modern war and terrorism, conventional, chemical, biological, nuclear.

All this is to be expected. Successful science produces knowledge, which facilitates the development of technology, both of which enormously increase our power to act. It is to be expected that this power will often be used beneficially (as it has been used), to cure disease, feed people, and in general enhance the quality of human life. But it is also to be expected, in the absence of wisdom, that such an abrupt, massive increase in power will be used to cause harm, whether unintentionally, as in the case (initially at least) of environmental damage, or intentionally, as in war and terror.

Before the advent of modern science, lack of wisdom did not matter too much; we lacked the means to do too much damage to ourselves and the planet. But now, in possession of unprecedented

powers bequeathed to us by science, lack of wisdom has become a menace. The crucial question becomes: How can we learn to become wiser?

The answer is staring us in the face. And yet it is one that almost everyone overlooks. Modern science has met with astonishing success in improving our knowledge of the natural world. It is this very success that is the cause of our current problems. But instead of merely blaming science for our troubles, as some are inclined to do, we need, rather, to try to learn from the success of science. We need to learn from the manner in which science makes progress towards greater knowledge how we can make social progress towards greater wisdom.

This is not a new idea. It goes back to the Enlightenment of the 18th century, especially the French Enlightenment. Voltaire, Diderot, Condorcet and the other *philosophes* of the Enlightenment had the profoundly important idea that it might be possible to learn from scientific progress how to achieve social progress towards an enlightened world. They did not just have the idea: they did everything they could to put the idea into practice in their lives. They fought dictatorial power, superstition, and injustice with weapons no more lethal than those of argument and wit. They gave their support to the virtues of tolerance, openness to doubt, readiness to learn from criticism and from experience. Courageously and energetically they laboured to promote reason and enlightenment in personal and social life. And in doing so they created, in a sense, the modern world, with all its glories and disasters.

The *philosophes* of the Enlightenment had their hearts in the right place. But in developing the basic Enlightenment idea intellectually the *philosophes*, unfortunately, blundered. They botched the job. And it is this that we are suffering from today. The *philosophers* thought that the proper way to implement the Enlightenment Programme of learning from scientific progress how to achieve social progress towards an enlightened world is to develop the social sciences alongside the natural sciences. If it is important to acquire knowledge of natural phenomena to better the lot of mankind, as Francis Bacon had insisted, then (so, in effect, the *philosophes* thought) it must be even more important to acquire

knowledge of social phenomena. First, knowledge must be acquired; then it can be applied to help solve social problems. They thus set about creating and developing the social sciences: economics, psychology, anthropology, history, sociology, political science.

This traditional version of the Enlightenment Programme, despite being damagingly defective, was immensely influential. It was developed throughout the 19th century, by men such as Saint-Simon, Comte, Marx, Mill and many others, and was built into the intellectual-institutional structure of academic inquiry in the first part of the 20th century with the creation of departments of the social sciences in universities all over the world.

Academic inquiry today, devoted primarily to the pursuit of knowledge and technological know-how, is the outcome of two past revolutions: the scientific revolution of the 16th and 17th centuries which led to the development of modern natural science, and the later profoundly important but very seriously defective Enlightenment revolution. It is this situation which calls for the urgent need to bring about a *third* revolution to put right the structural defects we have inherited from the Enlightenment.

But what, it may be asked, is wrong with the traditional Enlightenment Programme?

Almost everything. In order to implement properly the basic Enlightenment idea of learning from scientific progress how to achieve social progress towards a civilized world, it is essential to get the following three things right.

1. The progress-achieving methods of science need to be correctly identified.

2. These methods need to be correctly generalized so that they become fruitfully applicable to any worthwhile, problematic human endeavour, whatever the aims may be, and not just applicable to the one endeavour of acquiring knowledge.

3. The correctly generalized progress-achieving methods then need to be exploited correctly in the great human endeavour of trying to make social progress towards an enlightened, wise world.

Unfortunately, the *philosophes* of the Enlightenment got all three points wrong. And as a result these blunders, undetected and uncorrected, are built into the intellectual-institutional structure of academia as it exists today.

First, the *philosophes* failed to capture correctly the progress-achieving methods of natural science. From D'Alembert in the 18th century to Popper in the 20th, the widely held view, amongst both scientists and philosophers, has been (and continues to be) that science proceeds by assessing theories impartially in the light of evidence, *no permanent assumption being accepted by science about the universe independently of evidence*. But this standard empiricist view is untenable. If taken literally, it would instantly bring science to a standstill. For, given any accepted scientific theory, T, Newtonian theory say, or quantum theory, endlessly many rivals can be concocted which agree with T about observed phenomena but disagree arbitrarily about some unobserved phenomena. Science would be drowned in an ocean of such empirically successful rival theories if empirical considerations *alone* determined which theories are accepted, which rejected.

In practice, these rivals are excluded because they are disastrously disunified. *Two* considerations govern acceptance of theories in science: empirical success and unity. But in persistently accepting unified theories, to the extent of rejecting disunified rivals that are just as, or even more, empirically successful, science makes a big persistent assumption about the universe. Science assumes that the universe is such that all disunified theories are false. The universe has some kind of unified dynamic structure. It is physically comprehensible in the sense that explanations for phenomena exist to be discovered.

But this untestable (and thus metaphysical) assumption that the universe is comprehensible is profoundly problematic. How can we possibly *know* that the universe is comprehensible? Science is obliged to assume, but does not know, that the universe is comprehensible. Much less does it know that the universe is comprehensible in this or that way. A glance at the history of physics reveals that ideas about how the universe may be comprehensible have changed dramatically over time. In the 17th century there was the idea that the universe consists of corpuscles,

minute billiard balls, which interact only by contact. This gave way to the idea that the universe consists of point-particles surrounded by rigid, spherically symmetrical fields of force, which in turn gave way to the idea that there is one unified self-interacting field, varying smoothly throughout space and time. Nowadays we have the idea that everything is made up of minute quantum strings embedded in ten or eleven dimensions of space-time. Some kind of assumption along these lines must be made but, given the historical record, and given that any such assumption concerns the ultimate nature of the universe, that of which we are most ignorant, it is only reasonable to conclude that it is almost bound to be false.

The way to overcome this fundamental dilemma, inherent in the scientific enterprise, is to construe science as making a hierarchy of metaphysical assumptions concerning the comprehensibility and knowability of the universe, these assumptions asserting less and less as one goes up the hierarchy, and thus becoming more and more likely to be true. In this way a framework of relatively insubstantial, unproblematic, fixed assumptions and associated methods is created within which much more substantial and problematic assumptions and associated methods can be changed, and indeed improved, as scientific knowledge improves. Put another way, a framework of relatively unspecific, unproblematic, fixed *aims* and methods is created within which much more specific and problematic aims and methods evolve as scientific knowledge evolves. (A basic aim of science is to discover in what precise way the universe is comprehensible, this aim evolving as assumptions about comprehensibility evolve.) There is positive feedback between improving knowledge, and improving aims-and-methods, improving knowledge-about-how-to-improve-knowledge. This is the nub of scientific rationality, the methodological key to the unprecedented success of science. Science adapts its nature to what it discovers about the nature of the universe. For a detailed exposition and defence of this hierarchical, aim-oriented empiricist conception of science see my *The Comprehensibility of the Universe* (Oxford University Press, 1998, paperback 2003); see also my *The Human World in the Physical Universe* (Rowman and Littlefield, 2001) chapter 3 and

appendix 3, and *Is Science Neurotic?* (Imperial College Press, December 2004), chapters 1 and 2.

So much for the first blunder of the Enlightenment.

Second, having failed to identify the methods of science correctly, the *philosophes* naturally failed to generalize these methods properly. They failed to appreciate that the idea of representing the problematic aims (and associated methods) of science in the form of a hierarchy can be generalized and applied fruitfully to other worthwhile enterprises besides science. Many other enterprises have problematic aims; these would benefit from employing a hierarchical methodology, generalized from that of science, thus making it possible to improve aims and methods as the enterprise proceeds. There is the hope that, in this way, some of the astonishing success of science might be exported into other worthwhile human endeavours, with aims quite different from those of science.

Third, and most disastrously of all, the *philosophes* failed completely to try to apply such generalized progress-achieving methods to the immense, and profoundly problematic enterprise of making social progress towards an enlightened, wise world. The aim of such an enterprise is notoriously problematic. For all sorts of reasons, what constitutes a good world, an enlightened, wise or civilized world, attainable and genuinely desirable, must be inherently and permanently problematic. Here, above all, it is essential to employ the generalized version of the hierarchical, progress-achieving methods of science, designed specifically to facilitate progress when basic aims are problematic.

Properly implemented, in short, the Enlightenment idea of learning from scientific progress how to achieve social progress towards an enlightened world would involve developing social inquiry as social *methodology*, or social *philosophy*, not primarily as social *science*. A basic task would be to get into personal and social life, and into other institutions besides that of science – into government, industry, agriculture, commerce, the media, law, education, international relations – hierarchical, progress-achieving methods (designed to improve problematic aims) arrived at by generalizing the methods of science. A basic task for academic inquiry as a whole would be to help humanity learn how to resolve

its conflicts and problems of living in more just, cooperatively rational ways than at present. This task would be intellectually more fundamental than the scientific task of acquiring knowledge. Social inquiry would be intellectually more fundamental than physics. Academia would be a kind of people's civil service, doing openly for the public what actual civil services are supposed to do in secret for governments. Academia would have just sufficient power (but no more) to retain its independence from government, industry, the press, public opinion, and other centres of power and influence in the social world. It would seek to learn from, educate, and argue with the great social world beyond, but would not dictate. Academic thought would be pursued as a specialized, subordinate part of what is really important and fundamental: the thinking that goes on, individually, socially and institutionally, in the social world, guiding individual, social and institutional actions and life. The fundamental intellectual and humanitarian aim of inquiry would be to help humanity acquire wisdom – wisdom being the capacity to realize (apprehend and create) what is of value in life, for oneself and others, wisdom thus including knowledge and technological know-how but much else besides.

One important consequence flows from the point that the basic aim of inquiry would be to help us discover what is of value, namely that our feelings and desires would have a vital rational role to play within the intellectual domain of inquiry. If we are to discover for ourselves what is of value, then we must attend to our feelings and desires. But not everything that feels good is good, and not everything that we desire is desirable. Rationality requires that feelings and desires take fact, knowledge and logic into account, just as it requires that priorities for scientific research take feelings and desires into account. In insisting on this kind of interplay between feelings and desires on the one hand, knowledge and understanding on the other, the conception of inquiry that we are considering resolves the conflict between Rationalism and Romanticism, and helps us to acquire what we need if we are to contribute to building civilization: mindful hearts and heartfelt minds.

Another outcome of getting into social and institutional life the kind of aim-evolving, hierarchical methodology indicated above, generalized from science, is that it becomes possible for us to develop and assess rival philosophies of life as a part of social life, somewhat as theories are developed and assessed within science. Such a hierarchical methodology

“provides a framework within which diverse philosophies of value – diverse religions, political and moral views – may be cooperatively assessed and tested against the experience of personal and social life. There is the possibility of cooperatively and progressively improving such *philosophies of life* (views about what is of value in life and how it is to be achieved) much as *theories* are cooperatively and progressively improved in science. In science diverse universal theories are critically assessed with respect to each other, and with respect to *experience* (observational and experimental results). In a somewhat analogous way, diverse philosophies of life may be critically assessed with respect to each other, and with respect to *experience* – what we do, achieve, fail to achieve, enjoy and suffer – the aim being so to improve philosophies of life (and more specific philosophies of more specific enterprises within life such as government, education or art) that they offer greater help with the realization of value in life” (See my *From Knowledge to Wisdom*, 1984a, p. 254 or 2007, p.275).

All in all, if the Enlightenment revolution had been carried through properly, the three steps indicated above being correctly implemented, the outcome would have been a kind of academic inquiry very different from what we have at present. We would possess what we so urgently need, and at present so dangerously and destructively lack, institutions of learning well-designed from the standpoint of helping us create a better, a wiser world. Here, to conclude, is a summary of the changes that need to be made to science, and to academic inquiry more generally, to put right the blunders we have inherited from the Enlightenment, thus creating a kind of inquiry rationally designed to help humanity

realize what is genuinely of value, actually and potentially, in existence.

1. There needs to be a change in the basic intellectual *aim* of inquiry, from the growth of knowledge to the growth of wisdom — wisdom being taken to be the capacity to realize what is of value in life, for oneself and others, and thus including knowledge, understanding and technological know-how.
2. There needs to be a change in the nature of academic *problems*, so that problems of living are included, as well as problems of knowledge. Furthermore, problems of living need to be treated as intellectually more fundamental than problems of knowledge.
3. There needs to be a change in the nature of academic *ideas*, so that proposals for action are included as well as claims to knowledge. Furthermore, proposals for action need to be treated as intellectually more fundamental than claims to knowledge.
4. There needs to be a change in what constitutes intellectual *progress*, so that progress-in-ideas-relevant-to-achieving-a-more-civilized-world is included as well as progress in knowledge, the former being indeed intellectually fundamental.
5. There needs to be a change in the idea as to where inquiry, at its most fundamental, is located. It is not esoteric theoretical physics, but rather the thinking we engage in as we seek to achieve what is of value in life.
6. There needs to be a dramatic change in the nature of social inquiry (reflecting points 1 to 5). Economics, politics, sociology, and so on, are not, fundamentally, *sciences*, and do not, fundamentally, have the task of improving knowledge about social phenomena. Instead, their task is threefold. First, it is to articulate problems of living, and propose and critically assess possible solutions, possible actions or policies, from the standpoint of their capacity, if implemented, to promote wiser ways of living. Second, it is to promote such cooperatively rational tackling of problems of living throughout the social

world. And third, at a more basic and long-term level, it is to help build the hierarchical structure of aims and methods of aim-oriented rationality into personal, institutional and global life, thus creating frameworks within which progressive improvement of personal and social life aims-and-methods becomes possible. These three tasks are undertaken in order to promote cooperative tackling of problems of living — but also in order to enhance empathic or “personalistic” understanding between people as something of value in its own right. Acquiring knowledge of social phenomena is a subordinate activity, engaged in to facilitate the above three fundamental pursuits.

7. Natural science needs to change, so that it includes at least three levels of discussion: evidence, theory, and research aims. Discussion of aims needs to bring together scientific, metaphysical and evaluative consideration in an attempt to discover the most desirable and realizable research aims.
8. There needs to be a dramatic change in the relationship between social inquiry and natural science, so that social inquiry becomes intellectually more fundamental from the standpoint of tackling problems of living, promoting wisdom.
9. The way in which academic inquiry as a whole is related to the rest of the human world needs to change dramatically. Instead of being intellectually dissociated from the rest of society, academic inquiry needs to be communicating with, learning from, teaching and arguing with the rest of society — in such a way as to promote cooperative rationality and social wisdom. Academia needs to have just sufficient power to retain its independence from the pressures of government, industry, the military, and public opinion, but no more. Academia becomes a kind of civil service for the public, doing openly and independently what actual civil services are supposed to do in secret for governments.
10. There needs to be a change in the role that political and religious ideas, works of art, expressions of feelings, desires and values have within rational inquiry. Instead of being excluded, they need to be explicitly included and critically assessed, as possible indications and revelations of what is of

value, and as unmasking of fraudulent values in satire and parody, vital ingredients of wisdom.

11. There need to be changes in education so that, for example, seminars devoted to the cooperative, imaginative and critical discussion of problems of living are at the heart of all education from five-year-olds onwards. Politics, which cannot be taught by knowledge-inquiry, becomes central to wisdom-inquiry, political creeds and actions being subjected to imaginative and critical scrutiny.
12. There need to be changes in the aims, priorities and character of pure science and scholarship, so that it is the curiosity, the seeing and searching, the knowing and understanding of individual persons that ultimately matters, the more impersonal, esoteric, purely intellectual aspects of science and scholarship being means to this end. Social inquiry needs to give intellectual priority to helping empathic understanding between people to flourish (as indicated in 6 above).
13. There need to be changes in the way mathematics is understood, pursued and taught. Mathematics is not a branch of knowledge at all. Rather, it is concerned to explore problematic *possibilities*, and to develop, systematize and unify problem-solving methods.
14. Literature needs to be put close to the heart of rational inquiry, in that it explores imaginatively our most profound problems of living and aids personalistic understanding in life by enhancing our ability to enter imaginatively into the problems and lives of others.
15. Philosophy needs to change so that it ceases to be just another specialized discipline and becomes instead that aspect of inquiry as a whole that is concerned with our most general and fundamental problems — those problems that cut across all disciplinary boundaries. Philosophy needs to become again what it was for Socrates: the attempt to devote reason to the growth of wisdom in life.

This is the revolution we need to bring about in our traditions and institutions of learning, if they are to be properly and rationally designed to help us learn how to make progress towards a wiser world.

Chapter Nine

Non-Empirical Requirements Scientific Theories Must Satisfy Simplicity, Unity, Explanation, Beauty

Abstract

A scientific theory, in order to be accepted as a part of theoretical scientific knowledge, must satisfy both empirical and non-empirical requirements, the latter having to do with simplicity, unity, explanatory character, symmetry, beauty. No satisfactory, generally accepted account of such non-empirical requirements has so far been given. Here, a proposal is put forward which, it is claimed, solves the problem. It suffices to solve the problem for theoretical physics, in the first instance at least. A physical theory, in order to satisfy non-empirical requirements of unity etc., must be such that the same laws govern all possible phenomena to which the theory applies. Eight increasingly demanding versions of this requirement are distinguished. Some implications for our understanding of science are indicated.

1 - The Problem

A scientific theory, in order to be accepted as a part of theoretical scientific knowledge, must be sufficiently:

- (1) empirically successful;
- (2) empirically contentful;
- (3) simple, unified, explanatory, beautiful, elegant, harmonious, non-*ad hoc*, conceptually coherent, invariant, symmetrical, organic, inwardly perfect, non-aberrant (all terms used in this context by scientists and philosophers of science).

It is important to note that this third non-empirical requirement plays a crucial role in science, especially in physics, to the extent, even, of persistently over-riding empirical requirements. Given

any accepted physical theory, T, however successful empirically, it will always be possible to concoct endlessly many more empirically successful theories, T₁, T₂, etc., if non-empirical requirements can be ignored. T will make endlessly many predictions concerning phenomena not yet observed. Rivals to T can be concocted by modifying T in ad hoc ways so that each rival makes a different prediction for some unobserved phenomenon. Then independently testable and corroborated hypotheses can be added to these rivals, the result being a series of theories, T₁, T₂, etc., which have all the empirical success of T, have excess empirical content over T, this excess content being empirically corroborated. T₁, T₂, etc., are thus empirically more successful than T. Furthermore, almost all accepted physical theories run into empirical difficulties for some phenomena and are, on the face of it, refuted. T₁, T₂, etc., can be further modified in an entirely ad hoc, arbitrary fashion, so that these theories predict correctly the phenomena that ostensibly refute T, so that T₁, T₂, etc. are, in addition empirically successful where T is refuted. In scientific practice, of course, these rivals to T, much more empirically successful than T, are never considered at all because of their failure to satisfy non-empirical requirements. The fact that such empirically more successful theories are persistently ignored because of their unacceptably ad hoc, complex, disunified character means that non-empirical considerations are persistently over-riding empirical considerations in physics.¹⁹¹ Non-empirical considerations thus play an irreplaceable and fundamental role in science.

But what is this mysterious non-empirical feature of simplicity, unity, etc., that any acceptable scientific theory must possess? This is the problem I set out to solve in this paper.

It deserves to be noted that this is an absolutely fundamental problem in the philosophy of science. The solution is required for (a) a specification of scientific method, and (b) the solution to the problem of induction. Both points are demonstrated by the point made above, namely that non-empirical considerations persistently

¹⁹¹ See Maxwell (1998, 47-56) for additional arguments in support of the point.

over-ride empirical considerations when it comes to the acceptance of scientific theories.

Non-empirical considerations can have a purely pragmatic role in science: in certain contexts, we choose one formulation over another, or even one theory over another, not because we judge our choice to be more likely to be true, but because it is such that the equations are easier to solve, it is easier to extract useful predictions from the choice we make. Here, I ignore such pragmatic considerations, at least initially, and concentrate exclusively on non-empirical requirements judged to be indicative of truth or knowledge (however fallibly).

The following seven aspects of the problem can be distinguished.

(1) The terminological problem: How can simplicity, unity (etc.) be significant notions, having methodological significance, when the question of whether a theory is simple or complex, unified or disunified, will depend crucially on how the theory in question is formulated? A change of formulation can turn a simple theory into a complex one, and vice versa.

(2) How can degrees of simplicity, unity (etc.) be assessed?

(3) How many *different* features of theories are involved? The plethora of terms used by scientists and philosophers of science in this context does not inspire confidence that people know what they are talking about.

(4) How can one do justice to the fact that conceptions of simplicity or unity evolve with evolving knowledge? Three of Newton's four rules of reasoning concern simplicity (Newton, 1962, 398-400), and yet Newton's notions are different from those of a modern physicist.

(5) How can one do justice to ambiguity of judgements concerning the relative simplicity or unity of theories? Thus Newton's theory of gravitation seems in one way much simpler than Einstein's, but in another way more complex, or at least less unified.

(6) How is persistent preference for simple or unified theories in science, even against the evidence, to be justified? This is, it should be noted, the problem of induction. Solve this, and the problem of induction is solved.

(7) What implications does the solution to these problems have for science itself?

By far the most serious item on this list is (1). I shall concentrate on (1), and at the end will make a few remarks about (2) to (7).

Richard Feynman has provided the following amusing illustration of problem (1): see (Feynman et al. 1965, 25-10 – 25-11). Consider an appallingly complex universe governed by 10^{10} quite different, distinct laws. Even in such a universe, the true "theory of everything" can be expressed in the dazzlingly simple, unified form: $A = 0$. Suppose the 10^{10} distinct laws of the universe are:

(1) $F = ma$; (2) $F = Gm_1m_2/d^2$; etc.

Let $A_1 = (F - ma)^2$, $A_2 = (F - Gm_1m_2/d^2)^2$, etc., for all 10^{10} distinct laws. Let

10^{10}

$A = \sum_{r=1} A_r$. The true "theory of everything" of this universe can

now be formulated as: $A = 0$. (This is true if and only if each $A_r = 0$.)

Most scientists and philosophers of science recognize that non-empirical considerations of simplicity, etc., play an important role in science, but no one has been able so far to solve the terminological problem (problem (1)). Weyl (1963, 155) remarked correctly that "The problem of simplicity is of central importance for the epistemology of the natural sciences". Einstein (1949, 23) recognized the problem but confessed that he was not "without more ado, and perhaps not at all" able to solve it. Jeffrey and Wrinch (1921) suggested that simplicity could be identified with paucity of adjustable constants in equations, but unfortunately number of constants can be changed by changes of formulation. Popper (1959, ch. VII) proposed that simplicity is falsifiability, but unfortunately falsifiability can always be increased by adding on independently testable hypotheses which, in general, will drastically decrease simplicity. (Popper's adjunct proposal, in terms of dimension, does not work either, and is in any case subservient to falsifiability.) More recently Friedman (1974), Kitcher (1981) and Watkins (1984) have sought to identify

simplicity or unity with structural, formal or axiomatic features, but these attempts fail (see Salmon, 1989; Maxwell, 1998, 65-8). More recently, McAllister (1996), Weber (1999), Schurz (1999), and Bartelborth (2002) have tackled the problem without success – but see the excellent paper by Maudlin (1996) on unification of theoretical physics. One author has even declared recently of “a general ‘theory’ of unification” that “no such account is ... possible” (Morrison, 2000, 1).

2 The Proposed Solution

Previous attempts at solving the problem have failed because of mistakes concerning two crucial preliminary points.

The first mistake is to formulate the problem, in the first instance, too generally as a problem about *scientific* theories. It is vital, in the first instance, to restrict the problem to fundamental, dynamical *physical* theories. Branches of the natural sciences are not independent of one another; they are interconnected. Biology presupposes chemistry, and even physics, chemistry presupposes physics, geology and astronomy presuppose physics, and phenomenological physics presupposes fundamental physics. All branches of natural science besides theoretical physics, in other words, are constrained by results from some more fundamental science which, in the end, can be traced back to theoretical physics. This is neither a pro- nor anti-reductionist thesis; it is just the simple observation that theories in non-physical branches of natural science are in the end, in general, exceptions aside, constrained by physics. Only in fundamental theoretical physics does the question of the nature of non-empirical constraints on theories arise in something like a naked, pure form. We must, in the first instance, restrict the problem to that of fundamental, dynamical *physical* theory.

The second mistake is to suppose that simplicity, unity, etc., is a feature of the *theory* itself, its axiomatic structure, its simplicity of formulation, its number of postulates, its characteristic pattern of derivations, its number of adjustable constants. But all this involves looking at entirely the wrong thing. What one needs to look at is not the theory itself, but at the world, or rather at *what the theory says about the world*, the *content* of the theory in other

words. At a stroke the worst aspect of the problem of what unity *is* vanishes. No longer does one face the *terminological* problem of unity – the problem of the formulation-dependent nature of unity. Suppose we have a given theory T, which is formulated in N different ways, some formulations exhibiting T as beautifully unified, others as horribly complex and disunified, but all formulations being interpreted in precisely the same way, so as to make precisely the same assertion about the world. If unity has to do exclusively with *content*, then *all these diverse formulations of T, having the same content, have precisely the same degree of unity*. The variability of apparent unity with varying formulations of one and the same theory, T, (given some specific interpretation), which poses such an insurmountable problem for traditional approaches to the problem, poses no problem whatsoever for the thesis that unity has to do with *content*. Variability of formulation of a theory which leaves its content unaffected is wholly irrelevant: the unity of the theory is unaffected.

But now we have a new problem: How is the unity of the *content* of a theory to be assessed? What exactly does it mean to assert that a dynamical physical theory has a unified content?

What it means is that the theory has *the same* content throughout the range of possible phenomena to which the theory applies. Unity, in other words, means that there is just *one* content throughout the range of possible phenomena to which the theory applies. If the theory postulates *different* contents, *different* laws, for different ranges of possible phenomena, then the theory is *disunified*, and the more such *different* contents there are so the more *disunified* the theory is. Thus “unity” means “one”, and “disunity” means “more than one”, the disunity becoming worse and worse as the number of different contents goes up, from two to three to four, and so on. Not only does this enable us to distinguish between “unified” and “disunified” theories; it enables us to assign “degrees of unity” to theories, or to partially order theories with respect to their degree of unity.¹⁹²

¹⁹² If the theory is formulated as a set of differential equations, then what is invariant throughout the possible phenomena to which the theory applies is what is asserted by the physically interpreted set

To give an elementary example, Newton's theory of gravitation, $F = GM_1M_2/d^2$ is unified in that what the theory asserts is *the same* throughout all possible phenomena to which it applies (all bodies of all possible masses, constitution, shape, relative velocity, distance apart, at all times and places). An aberrant version of this theory, which asserts that $F = GM_1M_2/d^2$ for times $t \leq t_0$, where t_0 is some definite time, and $F = GM_1M_2/d^3$ for times $t > t_0$, is disunified because what the theory asserts is *not* the same throughout the range of possible phenomena to which the theory applies.

Note that special terminology could be introduced to make Newtonian theory look *disunified*, and the aberrant version of Newtonian theory look *unified*. All we need do is interpret " d^N " to mean " d^N if $t \leq t_0$ and d^{N+1} if $t > t_0$ ". In terms of this (admittedly somewhat bizarre) terminology, the aberrant theory has the form " $F = GM_1M_2/d^2$ ", and Newtonian theory has the "aberrant" form " $F = GM_1M_2/d^2$ for times $t \leq t_0$ and $F = GM_1M_2/d$ for times $t > t_0$ ". But this mere *terminological* reversal of aberrance or disunity does not affect the *content* of the two theories: the content of Newtonian theory remains unified, and the content of the aberrant version (which looks unified) remains disunified.

This almost suffices to solve the problem. A little more needs to be added, however, because in practice in physics assessments of degrees of unity are somewhat more complex than I have indicated so far because of the following consideration. In assessing the extent to which a theory is disunified we may need to consider *how* different, or *in what way* different, one from another, the different contents of a theory are. A theory that postulates different laws at different times and places is disunified in a much more serious way than a theory which postulates the same laws at all times and places, but also postulates that distinct kinds of physical particle exist, with different dynamical properties, such as charge or mass.

of differential equations. Laws specifying precisely how diverse physical states evolve in space and time may be quite diverse in character: what matters is that they are all solutions of the same set of differential equations.

This second theory still postulates *different* laws for different ranges of phenomena: laws of one kind for possible physical systems consisting of one kind of particle, and slightly different laws for possible physical systems consisting of another kind of particle. But this second kind of difference in content is much less serious than the first kind (which involves different laws at different times and places).

What this means is that there are different *kinds* of disunity, different *dimensions* of disunity, as one might say, some more serious than others, but all facets of the same basic idea. We can, I suggest, distinguish at least eight different facets of disunity, as follows.

Any dynamical physical theory, T, can be regarded as specifying an abstract space, S, of possible physical states to which the theory applies, a distinct physical state corresponding to each distinct point in S. (S might be a set of such spaces.) For unity, we require that T asserts that the same dynamical laws apply throughout S, governing the evolution of the physical state immediately before and after the instant in question. If T postulates N distinct dynamical laws in N distinct regions of S, then T has disunity of degree N. The eight different kinds of disunity can be characterized like this.

(1) T divides spacetime up into N distinct regions, $R_1...R_N$, and asserts that the laws governing the evolution of phenomena are the same for all spacetime regions within each R-region, but are different within different R-regions. Example: the aberrant version of Newtonian theory (NT) indicated above.

(2) T postulates that, for distinct ranges of physical variables (other than position and time), such as mass or relative velocity, in distinct regions, $R_1...R_N$ of the space of all possible phenomena, distinct dynamical laws obtain. Example: T asserts that everything occurs as NT asserts, except for the case of any two solid gold spheres, each having a mass of between one and two thousand tons, moving in otherwise empty space up to a mile apart, in which case the spheres attract each other by means of an inverse cube law of gravitation. Here, $N = 2$ in a type (2) way.

(3) In addition to postulating non-unique physical entities (such as particles), or entities unique but not spatially restricted (such as

fields), T postulates, in an arbitrary fashion, $N - 1$ distinct, unique, spatially localized objects, each with its own distinct, unique dynamic properties. Example: T asserts that everything occurs as NT asserts, except there is one object in the universe, of mass 8 tons, such that, for any matter up to 8 miles from the centre of mass of this object, gravitation is a repulsive rather than attractive force. The object only interacts by means of gravitation. Here, $N = 2$, in a type (3) way.

(4) T postulates physical entities interacting by means of N distinct forces, different forces affecting different entities, and being specified by different force laws. (In this case one would require one force to be universal so that the universe does not fall into distinct parts that do not interact with one another.) Example: T postulates particles that interact by means of Newtonian gravitation; some of these also interact by means of an electrostatic force $F = Kq_1q_2/d^2$, this force being attractive if q_1 and q_2 are oppositely charged, otherwise being repulsive, the force being much stronger than gravitation. Here, $N = 2$ in a type (4) way.

(5) T postulates N different kinds of physical entity,¹⁹³ differing with respect to some dynamic property, such as value of mass or charge, but otherwise interacting by means of the same force. Example: T postulates particles that interact by means of Newtonian gravitation, there being three kinds of particles, of mass m , $2m$ and $3m$. Here, $N = 3$ in a type (5) way.

(6) Consider a theory, T, that postulates N distinct kinds of entity (e.g. particles or fields), but these N entities can be regarded as arising because T exhibits some symmetry (in the way that the electric and magnetic fields of classical electromagnetism can be regarded as arising because of the symmetry of Lorentz invariance, or the eight gluons of chromodynamics can be regarded as arising as a result of the local gauge symmetry of $SU(3)$). If the symmetry

³ Counting entities is rendered a little less ambiguous if a system of M particles is counted as (a somewhat peculiar) field. This means that M particles all of the same kind (i.e. with the same dynamic properties) is counted as one entity. In the text I continue to adopt the convention that M particles all the same dynamically represents one kind of entity, rather than one entity.

group, G , is not a direct product of subgroups, we can declare that T is fully unified; if G is a direct product of subgroups, T lacks full unity; and if the N entities are such that they cannot be regarded as arising as a result of some symmetry of T , with some group structure G , then T is disunified.¹⁹⁴

(7) If (apparent) disunity of there being N distinct kinds of particle or distinct fields has emerged as a result of cosmic spontaneous symmetry-breaking events, there being manifest unity before these occurred, then the relevant theory, T , is unified. If current (apparent) disunity has not emerged from unity in this way, as a result of spontaneous symmetry-breaking, then the relevant theory, T , is disunified. Example: Weinberg's and Salam's electroweak theory, according to which at very high energies, such as those that existed soon after the big bang, the electroweak force has the form of two forces, one with three associated massless particles, two charged, W^- and W^+ , and one neutral, W^0 , and the other with one neutral massless particle, V^0 . According to the theory, the two neutral particles, W^0 and V^0 , are intermingled in two different ways, to form two new, neutral particles, the photon, γ , and another neutral massless particle, Z^0 . As energy decreases, the W^+ , W^- and Z^0 particles acquire mass, due to the mechanism known as spontaneous symmetry-breaking (involving the hypothetical Higgs particle), while the photon, γ , retains its zero mass. This theory unifies the weak and electromagnetic forces as a result of exhibiting the symmetry of local gauge invariance; this unification is only partial, however, because the symmetry group is a direct product of two groups, $U(1)$ associated with V^0 , and $SU(2)$ associated with W^-, W^+ and W^0 .¹⁹⁵

(8) According to GR, Newton's force of gravitation is merely an aspect of the curvature of spacetime. As a result of a change in our

¹⁹⁴ For accounts of the locally gauge invariant structure of quantum field theories see: Moriyasu (1983), Aitchison and Hey (1982: part III), and Griffiths (1987, ch. 11). For introductory accounts of group theory as it arises in physics see Isham (1989) or Jones (1990).

¹⁹⁵ For accounts of spontaneous symmetry breaking see Moriyasu (1983), Mandl and Shaw (1984), Griffiths (1987, ch. 11).

ideas about the nature of spacetime, so that its geometric properties become dynamic, a physical force disappears, or becomes unified with spacetime. This suggests the following requirement for unity: spacetime on the one hand, and physical particles-and-forces on the other, must be unified into a single self-interacting entity, U. If T postulates spacetime and physical "particles-and-forces" as two fundamentally distinct kinds of entity, then T is not unified in this respect. Example: one might imagine that the quantization of spacetime leads to the appearance of particles and forces as only apparently distinct from empty space-time. Here, $N = 1$ in a type (8) way: there is just the one self-interacting entity, empty spacetime.

For unity, in each case, as I have said, we require $N = 1$. As we go from (1) to (5), the requirements for unity are intended to be accumulative: each presupposes that $N = 1$ for previous requirements. As far as (6) and (7) are concerned, if there are N distinct kinds of entity which are not unified by a symmetry, whether broken or not, then the degree of disunity is the same as that for (4) and (5), depending on whether there are N distinct forces, or one force but N distinct kinds of entity between which the force acts.

(8) does not introduce a new kind of unity, but introduces, rather, a new, more severe way of counting different kinds of entity. (1) to (7) require, for unity, that there is one kind of self-interacting physical entity evolving in a distinct spacetime, the way this entity evolves being specified, of course, by a consistent physical theory. According to (1) to (7), even though there are, in a sense, two kinds of entity, matter (or particles-and-forces) on the one hand, and spacetime on the other, nevertheless $N = 1$. According to (8), this would yield $N = 2$. For $N = 1$, (8) requires that matter and spacetime are no more than aspects of one basic entity (unified by means of a spontaneously broken symmetry, perhaps).

As we go from (1) to (8), then, requirements for unity become increasingly demanding, with (6) and (7) being at least as demanding as (4) and (5), as explained above.

(1) to (8) may seem very different requirements for unity. In fact they all exemplify the same basic idea: disunity arises when

different dynamical laws govern the evolution of physical states in different regions of the space, S, of all possible physical states. For example, if a theory postulates more than one force, or kind of particle, not unified by symmetry, then in different regions of S different force laws will operate. If (8) is not satisfied, there is a region of S where only empty space exists, the laws being merely those which specify the nature of empty space or spacetime. The eight distinct facets of unity, (1) to (8) arise, as I have said, because of the eight *different* ways in which content can vary from one region of S to another.¹⁹⁶

3 - Objections

It may be objected that we never encounter the naked content of a theory, formulation free; we only encounter theories given some formulation. How, then, can we judge whether the *content* does or does not vary through the space S? The answer is that theories are not natural objects we stumble across; *we* formulate theories, and it is for us to ensure, granted we want our theories to be unified, that the content does not change as we move through S. We can arrange, however, that formulation matches content by ensuring that the terminology, the concepts, we use to formulate a theory do not surreptitiously change as we move through S. Given invariant concepts, if the form of the theory is also invariant throughout S, its content will be too. But if, for example, we surreptitiously change our units of length as we move through space, then a theory whose content is spatially invariant will change its form with changes of spatial position (a point which will be taken up again below).

It may be objected that, given any theory, however unified, special regularities will always arise in restricted regions of S, which means disunity. Whether or not the theory is unified is, at best, ambiguous. Thus, given NT, in some regions of S there will be solar systems with planets that rotate in the same direction and conform to Bode's law, whereas in other regions of S these regularities or "laws" will be violated. The answer is to distinguish

¹⁹⁶ This account of unity radically simplifies and improves on the account given in Maxwell (1998, chs. 3 and 4).

sharply between accidental and law-like regularities; only the latter are relevant for the assessment of unity. But how is this distinction to be made? The answer is to adopt a suggestion made elsewhere (Maxwell, 1968; 1998, 141-55) that physical laws are true iff corresponding physical dynamical (or necessitating) properties exist. According to this suggestion, Newton's law of gravitation can be interpreted as attributing the dynamical property of Newtonian gravitational charge to massive objects. Objects that have this property of necessity obey Newton's law of gravitation. (The empirical content of NT, on this interpretation, is concentrated in the factual assertion: all massive objects possess Newtonian gravitational charge equal to their mass.) If no such property corresponds to a true regularity, then it is merely a true accidental regularity, and not a true law. For unity we require that dynamical properties remain the same throughout S; the regularities of (some) solar systems, mentioned above, are not relevant because these regularities are not law-like, and no dynamical property exists corresponding to them.¹⁹⁷

It may be objected that physical systems which possess symmetries that are also symmetries of the theory which determines their evolution, will evolve in accordance with a simplified version of the theory. Thus systems consisting of two spheres equal in every way rotating in a fixed circle about their centre of mass obey a simplified version of the dynamical laws of NT. This means there are regions of S where the dynamical laws are especially simple, and thus different from other regions. Does this mean the theory is correspondingly disunified? The answer is No. We need, again, to consider dynamical properties corresponding to dynamical laws. In the example just considered, if NT is true (interpreted essentialistically) then the spheres in question possess gravitational charge just like all other massive objects. It is just that, in the case of the systems possessing some rotational symmetry, the full, rich implications of the dynamical property of gravitational charge is not made manifest.

¹⁹⁷ I am grateful to Jos Uffink for drawing my attention to the two objections just discussed.

It may be objected that we may not know whether two formulations of a theory are just that, two formulations with the same physical content, or two distinct theories with distinct contents. Heisenberg's and Schroedinger's distinct formulations of quantum theory might be an example. This is correct but beside the point. The terminological problem arises when we reformulate a given theory, T , in a variety of ways, some simple and unified, some horribly complex and disunified, but we do this in such a way as to ensure quite specifically that the different formulations have precisely the same content, make precisely the same assertions about the world, this being something that we can always do. The solution to the problem proposed above is not in any way undermined by the fact that it sometimes happens that we do not know whether two formulations of a theory have the same or different contents. Nor is the distinction between form and content undermined: form has to do with what we write down on paper, content with what is being asserted. That we sometimes do not know whether difference of formulation ensures difference of content does not in the least undermine the distinction between form and content. It deserves to be noted, in addition, that one and the same formulation of a theory may be interpreted in more than one way, and thus may have different contents associated with it – a point which, again, does not undermine the theory presented here.

It may be objected the distinction between dynamical laws which do, and do not, remain the same throughout the space S cannot be maintained. Consider the two functions (1) $y = 3x$ for all x , and (2) $y = 3x$ for $x \leq 2$ and $y = 4x$ for $x > 2$. It is tempting to say that (1) remains the same as x changes, but (2) does not, since what (2) asserts changes at $x = 2$. But given the mathematical notion of function as a rule, (2) is just as good a function as (1) and, like (1), “remains the same” as x varies. Functions corresponding to physical theories are somewhat more elaborate than this, but the above point is not affected by that consideration: it seems that the very distinction between “remains the same” and “changes” as one moves through S collapses. Clearly, in order to meet this objection, functions corresponding to physical theories need to be restricted to a narrower notion of function than the

above standard mathematical one, if we are to be able to distinguish between functional relationships which do, and which do not, “remain the same” as values of variables change. We need to appeal to what may be called “invariant functions”, functions which specify some fixed set of mathematical operations to be performed on “x” (or its equivalent) to obtain “y” (or its equivalent). In the example just given, (1) is invariant, but (2) is not. (2) is made up of two truncated invariant functions, stuck together at $x = 2$. Functions that appear in theoretical physics are *analytic*; that is, they are repeatedly differentiable. Such functions have the remarkable property that from any small bit of the function, the whole function can be reconstructed uniquely, by a process called “analytic continuation”. All analytic functions are thus invariant. The latter notion is however a wider one, and theoretical physics might, one day, need to employ this wider notion explicitly, if space and time turn out to be discontinuous, and analytic functions have to be abandoned at a fundamental level.

A similar remark needs to be made about Goodman’s (1954) paradox concerning “grue” and “bleen”. Modifying the paradox slightly, an object is grue if it is green up to time t , blue after t ; it is bleen if it is blue up to time t , green after t . Sometimes it is held that there is perfect symmetry between blue and green, on the one hand, and grue and bleen on the other, especially as “emeralds are green” is equivalent to “emeralds are grue up to t , and bleen afterwards”. But this symmetry is merely terminological and, as we have already seen in connection with the aberrant version of Newton’s theory of gravitation, discussed in section 2, terminological symmetry does not mean there is symmetry of content. That there is not symmetry of content in the grue/bleen case can be demonstrated as follows. If emeralds are grue, a person convinced of this can determine whether t is future or past merely by looking at emeralds. But if emeralds are green, a person convinced of this cannot say whether t is in the future or past by just looking at emeralds. The content, the meaning, of grue and bleen contains an implicit reference to t in a way in which that of green and blue do not. Doubtless symmetry can be created by considering two possible worlds, ours and a Goodmanesque one

with special physics and/or physiology of vision so that grue emeralds do not appear to change at t , whereas green ones do. This, however, is to consider conditions quite different from those specified by Goodman. The crucial point to make, in any case, is that dynamical or physical properties, of the kind attributed to physical entities by physical theories (interpreted in a conjecturally essentialistic way), are like blue and green, and unlike grue and bleen, in not containing any implicit reference to specific times or places (or hypersurfaces of S that distinguish one region of S from another). Physical properties must be *invariant* in a sense that corresponds to the *invariance* of allowed functions in physics. The more general notion of property, which includes Goodmanesque properties, is excluded, just as the more general notion of function, which includes (2) above as an “unchanging” function, is excluded.

4 – Further Issues

What of the other aspects of the problem of non-empirical requirements in science mentioned in the introduction? I have space, here, only for staccato remarks concerning some of these further issues.

Most of the other terms used to refer to non-empirical requirements can be straightforwardly related to unity. We have seen that this is true of symmetry and invariance. Non-ad hoc, organic, inwardly perfect and non-aberrant can be interpreted as appealing to unity, and harmonious, beautiful and conceptually coherent can be interpreted as presupposing unity. A dynamical physical theory can be held to be explanatory in character to the extent that it is (1) empirically contentful and (2) unified.

Simplicity, however, is quite different. The simplicity of a theory can be interpreted as having to do, not with whether the *same* laws apply throughout the space S , but rather with the *nature* of the laws, granted that they are the same. Some laws are simpler than others. In order to overcome the objection that simplicity is formulation dependent it is essential, as in the case of unity, to interpret “simplicity” as applying to the *content* of theories, and not to their *formulation*, their *axiomatic structure*, etc. Theories can only, at best, be partially ordered with respect to degrees of

simplicity. Even when two theories are amenable to being assessed with respect to relative simplicity, there is always the problem that a change of variables may reverse the assessment. Let the two theories be (1) $y = x$ and (2) $y = x^2$. We judge (1) to be simpler than (2). Let $x^2 = z$. We now have (1) $y = \sqrt{z}$, and (2) $y = z$. Now (2) is simpler than (1). Assessment of relative simplicity of two theories may only be unambiguous when restrictions are placed on the form that physical variables can take, so that only linear transformations of the type $z = Ax + B$ (where A and B are constants) are permitted, for example. It is a further great success of the theory presented here that it succeeds in distinguishing sharply between these two aspects of physical theory, the *unity* and *simplicity* aspects, and succeeds in explicating both.¹⁹⁸

We can use these two notions to solve the problem of ambiguity of judgment concerning the relative non-empirical merits of Newton's and Einstein's theories of gravitation. Newton's theory is simpler, but Einstein's is more unified in that it eliminates gravitation as a force, and reformulates Newton's first law so that it becomes the assertion that bodies move along geodesics in curved space-time, curvature being caused by mass, or by stress-energy-density more generally. As theoretical physics draws closer to capturing the true theory of everything it is reasonable to expect that the totality of fundamental physical theory will become increasingly unified and complex.

So far I have stressed that terminological unity and simplicity are irrelevant when it comes to assessing unity and simplicity in a physically significant sense. In scientific practice, however, terminology is chosen so as to reflect physically significant unity and simplicity (Maxwell, 1998, 110-3). Thus if the content of a theory exhibits certain symmetries, terminology is chosen so that it too exhibits these symmetries, so that if the theory is invariant with respect to position or orientation in space, terminology is chosen which reflects this fact. Once a theory is formulated in such "physically appropriate" terminology (as it may be called), two

¹⁹⁸ For further discussion of simplicity, and how terminological simplicity can be related to unity, see Maxwell (1998, 110-3 and 157-9).

versions of symmetry operations arise as a result, “active” (which make changes to physical systems) and “passive” (which make corresponding changes to the description of unchanged physical systems). Granted that we formulate physical theories exclusively in such “physically appropriate” terminology, then terminological unity and simplicity comes to reflect physical unity and simplicity, and is thus, to that extent, physically significant.

What of the simplicity and unity of theories in sciences other than fundamental physics? Much needs to be said on this topic; the following brief remarks can serve only as pointers to a more adequate treatment. Solutions to the equations of fundamental physical theory, specifying precisely how increasingly complex physical systems evolve in space and time, rapidly become horrendously complex in character. In carrying out derivations, physicists invariably “simplify” results obtained by discarding variable quantities or higher order terms judged to be insignificant in the physical situations under considerations. Just this is done when NT is “derived” from Einstein’s theory, or Kepler’s and Galileo’s laws of motion are “derived” from NT.¹⁹⁹ The outcome is a range of more or less terminologically simple phenomenological laws of only approximate validity. But the simplicity is not, here, merely pragmatic, since such a law has been “approximately derived” from some fundamental physical theory formulated in a “physically appropriate” way, the “approximate derivation” showing what the range of applicability of the law is with what degree of accuracy. Even though such laws are incompatible with the fundamental physical theory from which they have been “approximately derived”, nevertheless what the “derivations” reveal is that pragmatic simplicity has been obtained by sacrificing strict derivability and precise empirical accuracy, there being nothing here to counter the underlying unity in nature postulated by fundamental physical theory (insofar as it does postulate this). Laws such as these are prevalent throughout phenomenological physics, astrophysics and parts of physical chemistry. Even where such “approximate derivations” cannot be

¹⁹⁹ For a discussion of such “approximate derivations”, the conclusion being strictly incompatible with the premises, see Maxwell (1988, 211-7).

carried through, for large parts of chemistry, and for biology, nevertheless, as I have already remarked, laws and theories of these sciences are constrained by fundamental physics, and must endeavor to be compatible with fundamental physics, at least in the qualified way just indicated in connection with phenomenological physics. Thus, much of the great explanatory power of Darwinian theory stems from the fact that it postulates mechanisms for evolution – random inheritable variation and natural selection – which are capable of designing living things able to pursue the goals of survival and reproductive success in their given environments, these mechanisms nevertheless being compatible with the purposeless cosmos depicted by physics. Biology must accord with physics in much more specific ways as well, in that the mechanisms of inheritance and development must accord with physics, and so too the multitude of processes that take place in living things.

What implications does the account of non-empirical requirements for theories, given here, have for science? How can justice be done to evolving non-empirical requirements? How is persistent preference for unified theories, even against the evidence, to be justified? I take these three problems together.

At the beginning of this paper I demonstrated that, in physics, theories that are unified, in senses (1) and (2) at least, are persistently chosen in preference to available, empirically more successful, but disunified theories. To proceed in this way is to make the permanent assumption that the phenomena under consideration are such that all theories of these phenomena that are disunified in senses (1) and (2) are false. If physicists persistently accepted theories that postulate atoms in preference to available, empirically more successful *field* theories, it would be clear that physicists are thereby assuming that all field theories are false. Just the same holds for the persistent rejection of empirically more successful disunified theories.

But rigour demands that assumptions that are substantial, influential, problematic and implicit need to be made explicit, so that they can be critically assessed, so that alternatives can be developed and considered, the hope being that in this way such assumptions can be improved. Thus rigour demands that science

makes explicit, and so criticizable and improvable, the substantial, problematic, influential and implicit assumption that the universe, or the phenomena, are such that all disunified theories are false. This assumption, M, can easily be shown to be metaphysical, as follows. Persistent acceptance of theories unified in ways (1) and (2) involves rejecting infinitely many empirically more successful disunified rivals, $T_1, T_2, \dots T_\infty$, because they clash with M. In effect, $M = \text{not}T_1$ and $\text{not}T_2$ and ...and $\text{not}T_\infty$. In order to verify M we would need to falsify all of T_1, T_2, \dots and T_∞ , but as there are infinitely many theories, this cannot be done. In order to falsify M we need to verify just one of T_1, T_2, \dots or T_∞ , but physical theories cannot be verified. Hence M, being neither verifiable nor falsifiable, is metaphysical. It is a permanent metaphysical assumption of science – permanent, at least, as long as all theories disunified in senses (1) and (2) are rejected whatever their empirical success might be.

At once the question arises: How is this assumption M to be critically assessed and, perhaps, improved? Elsewhere I have argued that once the metaphysical assumption implicit in persistent preference in science for unified theories are acknowledged, it becomes apparent that we need to adopt a new conception of science, which construes science as making a hierarchy of such assumptions, these assumptions asserting less and less as one goes up the hierarchy, and thus becoming more and more likely to be true (see Maxwell, 1998). These are assumptions about the knowability and comprehensibility of the universe. As we descend the hierarchy, assumptions become more substantial and specific, and much more likely to be false, and in need of revision. Revision is, however, kept as low down in the hierarchy as possible. Those physical theories are accepted which best accord with the evidence and the best available metaphysical assumption, B say, lowest down in the hierarchy. But B may itself be revised if a rival assumption, B^* , is developed which (a) is compatible with the assumption above it in the hierarchy, and (b) supports an empirical research programme that is more successful than the one supported by B. Relatively problematic assumptions high up in the hierarchy thus form a fixed framework within which much more

specific, problematic assumptions can be revised in the light of empirical success and failure. As knowledge improves, assumptions and associated methods improve as well; there is something like positive feedback between improving knowledge and improving knowledge-about-how-to-improve-knowledge, the methodological key to the success of modern science. Non-empirical requirements for theory acceptance, corresponding to metaphysical assumptions, improve with improving knowledge. Newton's requirements of simplicity evolve into the symmetry principles of modern physics. For a suggestion as to how acceptance of the hierarchy of metaphysical assumptions is to be justified, see Maxwell (1998, ch. 5).

Chapter Ten

Scientific Metaphysics

(Modified version of a paper posted on the Pittsburgh PhilSci Archive in 2004: <http://philsci-archive.pitt.edu/1674/> .)

1 – Introduction

The idea that science cannot proceed without the assumption that the universe is comprehensible in some way – the thesis that it is *physically* comprehensible being a more secure item of knowledge than any accepted physical theory – represents a profound and dramatic revolution in our whole conception of science. It is perhaps not surprising that, even though this idea dates back at least to 1974,²⁰⁰ and has been expounded and argued for in considerable detail on a number of occasions since,²⁰¹ nevertheless, at the time of writing (2011), few scientists or philosophers of science have taken note of, or responded to the arguments for, this revolutionary view.²⁰² Given the momentous consequences of the idea, and its neglect, it seems to me appropriate to reformulate, in

1. The date of my first publication arguing for aim-oriented empiricism: see Maxwell (1974).

2. See especially Maxwell (1998); see also Maxwell (1976a, ch. 6; 1984a, chs. 5 and 9; 1993a; 1997a; 1999a; 2000b; 2001a, ch. 3 and appendix 3; 2004; 2007, ch. 14).

3. Philosophers of science who have praised or criticized aim-oriented empiricism include Kneller (1978, pp. 80-87, 90-91); Harré (1986, pp. 26-32); Midgley (1986); Chakravartty (1999); Chart (2000); Juhl (2000); Shanks (2000); Smart (2000); Weinert (2000); McHenry (2000; 2009); Roush (2001); MacIntyre (2009); Muller (2004; 2008); Schiff (2005); Miller (2006, pp. 92-94); Rogers (2009); Vicente (2010); Pandit (2010). For my replies to criticisms, see Maxwell (2001c; 2006; 2009b; 2010b).

as careful and critical a way as I can, the central argument for aim-oriented empiricism. The argument developed in this chapter improves on the earlier versions, to be found in chapters 4 and 5.

I argue, first, that persistent acceptance of (more or less) *unified* fundamental dynamic theories in physics, even though endlessly many empirically more successful *disunified* rivals are always available, means that physics makes a persistent untestable (or metaphysical) assumption about the nature of the universe: it is such that some yet-to-be-discovered, more or less unified physical theory is true, and all seriously disunified theories are false. I then invoke the account of what it means to assert of a physical theory that it is unified, developed in the last chapter, to throw light on the question of what physics does, and ought to, assume in assuming that some more or less unified theory is true. This provides us with a way of classifying – indeed, of partially ordering – all metaphysical theses which assert that the universe possesses some kind of comprehensive, more or less disunified dynamic structure – the universe being more or less physically comprehensible, in other words.

At once the fundamental problem arises: How can physics choose between these infinitely many metaphysical theses – versions of *physicalism* as I call them? Two considerations drive us in opposite directions. On the one hand we ought to choose that version of physicalism most fruitful for promoting progress in theoretical physics, if true. On the other hand, we ought to choose that minimalist version of physicalism just sufficiently substantial to make physics possible, and thus least likely to be false.

I take these in turn. First, I spell out how the account of unity of the last chapter provides the means for assessing the relative fruitfulness of rival versions of physicalism, go on to specify that version of physicalism which is the most fruitful, given the history of physics up till today, and argue that physics should accept this version as its basic metaphysical assumption. Second, I consider the grounds for physics accepting the least substantial version of physicalism that makes physics possible. The problem of choosing between these conflicting considerations is solved by the hierarchical view of aim-oriented empiricism (AOE). This does justice to both apparently conflicting considerations – a strong

argument in favour of AOE. In conclusion I consider two versions of AOE, and indicate, briefly, how the circularity problem can be solved.

The title is intended to be provocative. “Scientific Metaphysics” sounds like a contradiction in terms in view of Popper’s well-known demarcation criterion which rules that metaphysical theses, being unfalsifiable, are not scientific. But of course Popper’s falsificationist conception of science, along with others, will be found to be defective precisely because of a failure to acknowledge the role that metaphysical assumptions play in science. Furthermore, as I have indicated, a framework will be developed which makes it possible to appraise (untestable) metaphysical theses empirically, in terms of their “empirical fruitfulness”, or fruitfulness for the empirical research programme of theoretical physics. For physics to be rigorous, it will be argued, it is essential that metaphysical theses are acknowledged as key components of theoretical knowledge in physics, and are appraised empirically in terms of their “empirical fruitfulness”. Once the conception of physics defended here is accepted, the title entirely loses its air of being self-contradictory.²⁰³

2 – Intellectual Rigour Requires that Metaphysical Presuppositions be made Explicit

Many views about science deny that science makes a substantial, persistent, metaphysical (i.e. untestable) assumption about the

4. It is widely appreciated that some metaphysical theses have influenced science in the context of discovery, in influencing scientists to try to develop certain sorts of theories, and to ignore others: see, for example, Watkins (1958), Popper (1959, p. 278). What is being argued here is very different. I argue that there are metaphysical theses, neither falsifiable nor verifiable, which are an integral part of theoretical scientific knowledge, more firmly established, indeed, than such highly corroborated theories as quantum theory and general relativity. Science is more rigorous if this is acknowledged rather than denied. All this differs dramatically even from Popper’s later views concerning the important role that metaphysical research programmes play in science: see Popper (1976a, sections 33 and 37; 1983, section 23; 1982a, sections 20-28). Popper held on to his demarcation criterion to the end, and never wavered from holding metaphysical theses, however scientifically fruitful, to be “unscientific”: for discussion of this point see ch. 4 and my 2012).

universe. This is true, for example, of logical positivism, inductivism, logical empiricism, hypothetico-deductivism, conventionalism, constructive empiricism, pragmatism, realism, Bayesianism, induction-to-the-best-explanationism, and the views of Popper, Kuhn and Lakatos. All these views, diverse as they are in other respects, accept a thesis that may be called standard empiricism (SE): in science, theories are accepted on the basis of empirical success and failure, and on the basis of simplicity, unity or explanatoriness, but *no substantial thesis about the world is accepted permanently by science, as a part of scientific knowledge, independently of empirical considerations.*²⁰⁴

The following argument shows however that SE is untenable.

Whenever a fundamental physical theory is accepted as a part of theoretical scientific knowledge there are always endlessly many rival theories which fit the available evidence just as well as the accepted theory. Consider, for example, Newtonian theory (NT). One rival theory asserts: everything occurs as NT asserts up till midnight tonight when, abruptly, an inverse cube law of gravitation comes into operation. A second rival asserts: everything occurs as NT asserts, except for the case of any two solid gold spheres, each having a mass of a thousand tons, moving in otherwise empty space up to a mile apart, in which case the spheres attract each other by means of an inverse cube law of gravitation. There is no limit to the number of rivals to NT that can be concocted in this way, each of which has all the predictive success of NT as far as observed phenomena are concerned but which makes different predictions for some as yet unobserved phenomena.²⁰⁵ Such theories can even be concocted which are

²⁰⁴. For discussion of the claim that Kuhn and Lakatos defend versions of SE see ch. 4 and my (1998, p. 40).

²⁰⁵. All the possible phenomena, predicted by any dynamical physical theory, T, may be represented by an imaginary "space", S, each point in S corresponding to a particular phenomenon, a particular kind of physical system evolving in time in the way predicted by T. In order to specify severely disunified rivals to T that fit all available evidence just as well as T does, all we need do is specify a region in S that consists of phenomena that have not been observed, and then replace the phenomena predicted by T with anything we care to think of. Given any T, there will always be infinitely many such disunified rivals to T. This point is inherent in Nelson Goodman's "new paradox of induction" (see Goodman, 1954),

more empirically successful than NT, by arbitrarily modifying NT, in just this entirely *ad hoc* fashion, so that the theories yield correct predictions where NT does not, as in the case of the orbit of Mercury for example (which very slightly conflicts with NT).²⁰⁶ And quite generally, given any accepted physical theory, T, there will always be endlessly many *ad hoc* rivals which meet with all the empirical success of T, make untested predictions that differ from T, are empirically successful where T is ostensibly refuted, and successfully predict phenomena about which T is silent (as a result of independently testable and corroborated hypotheses being added on).

As most physicists and philosophers of physics would accept, *two* criteria are employed in physics in deciding what theories to accept and reject: (1) empirical criteria, and (2) criteria that have to do with the simplicity, unity or explanatory character of the theories in question. (2) is absolutely indispensable, to such an extent that there are endlessly many theories empirically more successful than accepted theories, all of which are ignored because of their lack of unity.

Now comes the crucial point. In persistently accepting unifying theories (even though ostensibly refuted), and excluding infinitely many empirically more successful, unrefuted, disunified or aberrant rival theories, science in effect makes a big assumption about the nature of the universe, to the effect that it is such that some yet-to-be-discovered, more or less unified physical theory is true, and no seriously disunified theory is true²⁰⁷ however

although the kind of empirically successful disunified rivals considered by Goodman in his discussion of “grue” and “bleen” are but one kind of a number of kinds of disunified theories, as we shall see in section 3. There is a vast philosophical literature on the underdetermination of theory by evidence: for an excellent recent discussion, and reference to further literature, see Howson (2000): see especially ch. 1, pp. 30-4, 75-7, and ch. 5. See also Lipton (2004).

²⁰⁶. For a more detailed discussion of empirically successful *ad hoc* rivals to accepted theories, see Maxwell: (1974; 1993a; 1998, pp. 51-4).

8. Vicente (2010) claims that this argument cannot establish that physics assumes the *truth* of some more or less unified physical theory, and the *falsity* of all seriously disunified rival theories. It is however especially in the context of practical applications that physics requires predictions of accepted theories to be *true*, and clashing predictions of rival theories to be *false*. In contexts such as

empirically successful it may appear to be for a time. Furthermore, without some such big assumption as this, the empirical method of science collapses. Science would be drowned in an infinite ocean of empirically successful disunified theories.²⁰⁸

If scientists only accepted theories that postulate atoms, and persistently rejected theories that postulate different basic physical entities, such as fields – even though many field theories can easily be, and have been, formulated which are even more empirically successful than the atomic theories – the implications would surely be quite clear. Scientists would in effect be assuming that the world is made up of atoms, all other possibilities being ruled out. The atomic assumption would be built into the way the scientific community accepts and rejects theories – built into the implicit *methods* of the community, methods which include: reject all theories that postulate entities other than atoms, whatever their empirical success might be. The scientific community would accept the assumption: the universe is such that no non-atomic theory is true.

Just the same holds for a scientific community which rejects all disunified or aberrant rivals to accepted theories, even though these

bridge building, for example, we do indeed need to assume that relevant accepted laws will yield *true* predictions, and better empirically established but disunified rivals, that predict the bridge will collapse, are *false*. (Human lives are at stake.) In making such an assumption, against the evidence, we implicitly assume the truth of a metaphysical thesis concerning the unity of phenomena. For a more detailed rebuttal of Vicente's claim, see my (2010b, pp. 673-4).

9. It may be objected that the universe might have been genuinely disunified, so that physics could consist only of a great number of physical laws. In this case, it may be argued, physics could not be construed as making a metaphysical assumption about underlying unity. But even in this counterfactual situation, endlessly many very much more disunified but empirically more successful rival laws could easily be formulated: these would have to be rejected on non-empirical grounds, or physics would drown in an ocean of rival laws. The persistent rejection of such much more disunified but empirically more successful rivals would involve the methods of physics making an implicit metaphysical assumption, to the effect that nature is unified to some extent at least (all grossly disunified laws being false). It is necessary to make some such assumption, however disunified the totality of accepted laws may be – even if the assumption made is rather weak in character, in that only gross disunity is denied.

rivals would be even more empirically successful if they were considered. Such a community in effect makes the assumption: the universe is such that no disunified theory is true. Or rather, more accurately, such a community makes the assumption: “no disunified theory is true *which is not entailed by a true unified theory (plus, possibly, true relevant initial and boundary conditions)*”. (A true unified theory entails infinitely many approximate, true, disunified theories.) Let us call this assumption “physicalism”.

That physicalism is metaphysical can be shown as follows. Physicalism asserts “not T_1 and not T_2 and . . . and not T_∞ ”, where $T_1, T_2, . . . T_\infty$ are infinitely many disunified rivals to accepted physical theories. Physicalism cannot be empirically verified, because this would require that all of $T_1, T_2, . . . T_\infty$ are falsified, but as there are infinitely many of these theories, each requiring a different falsifying experiment, this cannot be done. Equally, physicalism cannot be falsified, as this requires the verification of at least one of $T_1, T_2, . . . T_\infty$, which cannot be done as physical theories cannot be verified empirically. Hence physicalism, being neither verifiable nor falsifiable, is metaphysical.

Thus in persistently rejecting empirically more successful but disunified rivals to accepted physical theories, science makes a persistent metaphysical assumption about the world, namely physicalism. Standard empiricism (SE), and all the above doctrines that include SE as a component, which hold science makes no persistent metaphysical assumption, are thus untenable.

Let us call the view that science presupposes physicalism “presuppositionism”. Presuppositionism is more rigorous than all the above versions of SE *entirely independent of any justification for accepting physicalism as a part of scientific knowledge* (that is in addition to the one given above). In saying this I am appealing to the following wholly uncontroversial requirement for rigour.

Principle of Intellectual Rigour (PIR) In order to be rigorous, it is necessary that assumptions that are substantial, influential and problematic be made explicit – so that they can be criticized, so

that alternatives may be developed and assessed, with the aim of improving the assumptions.²⁰⁹

All versions of SE fail to satisfy PIR in just the way in which presuppositionism does satisfy PIR. Presuppositionism makes the assumption of physicalism explicit (and so criticizable and, we may hope, improvable), while all versions of SE deny that science does make any such assumption as physicalism. Thus, quite independent of any claim to solve the problem of induction, presuppositionism is more rigorous, and thus more acceptable, than any of the above versions of SE. And this is the case *even though presuppositionism can provide no justification for accepting physicalism*. It is indeed above all when we have no reason whatsoever for supposing physicalism is true that it becomes all the more important to implement PIR, and make the probably false assumption of physicalism explicit, so that it can be critically assessed, so that alternatives can be considered, in the hope that a thesis nearer the truth can be discovered.

Why has this simple argument been ignored by the vast literature on the problem of induction (or underdetermination), referred to in note 6? Two factors are perhaps at work. First, accepting the argument involves acknowledging that science, as it is ordinarily understood (in terms of standard empiricism), lacks rigour. Our understanding of science, and even science itself, need to change if science is to become rigorous and make explicit, and critically assess, implicit, problematic metaphysical presuppositions. Philosophers, tackling the problem of induction, have perhaps been reluctant to take seriously that science, as ordinarily understood, is irrational and needs to be changed. Second, invoking a metaphysical presupposition of unity looks superficially like a well-known and hopelessly invalid approach to the problem of induction: justify scientific theory by an appeal to a metaphysical principle uniformity, and then justify this metaphysical principle by an appeal to the success of science. Aware of the vicious circularity of any such argument, philosophers have instinctively resisted considering the superficially similar, but actually very

²⁰⁹. See Maxwell: (1984a, p. 224; 1998, p. 21).

different, point that rigour requires that substantial, influential, problematic and implicit metaphysical presuppositions need to be made explicit – so that they can be critically assessed and, we may hope, improved.

Presuppositionism does not, perhaps, entirely solve the problem of induction – the problem of underdetermination – but it does, when further developed, transform that philosophical and scientifically sterile problem into the scientifically fruitful problem of developing and choosing the most fruitful metaphysics for physics, as we shall see below.

3 – Unity of Physical Theory

We have seen that physics only accepts theories that are unified, and this commits physics to presupposing physicalism. But what ought physicalism to be interpreted to assert, especially if physics is to comply with PIR? In order to answer this question we first need to solve the problem of what it means to assert of a physical theory that it is *unified*. This is a problem that has long resisted solution.²¹⁰ Even Einstein (1969, p. 23) confessed that he did not know how to solve the problem.

However, in the last chapter, section 2, we saw how the problem can be solved. Above, in the present chapter, I indicated ways in which theories can be *disunified*. The solution to the problem of what it is for a theory to be unified, spelled out in the previous chapter, in effect extends and develops the above remarks about disunity.

A dynamical physical theory is disunified if its content, what it asserts about the world, is *different*, from one region to another in the space of all possible phenomena to which the theory applies. If the content of the theory differs in N ways, throughout the space of phenomena predicted by the theory, then the theory is disunified to degree N. For unity, we require that $N = 1$.

But, as we saw in the last chapter, there is a refinement. There are *different* ways in which the content of a theory may differ, from one region in the space of possible phenomena to another,

11. See Salmon (1989) for a survey of attempts to solve the problem. See also Maxwell (1998, pp. 56-68).

some ways being more substantial, more serious as it were, than others. The most serious kind of difference, perhaps, is when the content of a theory differs in different space-time regions. A less serious kind of difference arises when a theory predicts no variation in dynamical laws in space-time, but predicts that there is more than one kind of force, or more than one kind of particle or field (there being one force, or one kind of particle or field in one sub-region of the space of all possible phenomena, and another force, or another kind of particle or field in another sub-region). In all, as we saw, there are *eight* different ways, (1) to (8), in which the content of a theory can differ, in different regions of the space of all possible phenomena that the theory predicts.

As we go from (1) to (8), the requirements for unity become increasingly demanding, with (6) and (7) being at least as demanding as (4) and (5), as explained in section 2 of chapter nine.²¹¹ It is important to appreciate, however, that (1) to (8) are all versions of the same basic idea that T is unified if and only if the content of T is the same throughout the range of possible phenomena to which it applies. When T is disunified, (1) to (8) specify *different* kinds of difference in the content of T in diverse regions of the space, S, of all possible phenomena to which T applies. Or, equivalently, (1) to (8) divide S into sub-regions in different ways, T having a different content in each sub-region. For (1), sub-regions contain physical systems in different locations in spacetime, the content of T being different in different spacetime locations. For (2), sub-regions contain physical systems with different values of physical variables such as mass or relative velocity. For (3), sub-regions contain systems with different dynamically unique objects. For (4), sub-regions contain systems

12. The account of theoretical unity given here simplifies the account given in (Maxwell 1998, chs. 3 and 4), where unity is explicated as "exemplifying physicalism", where physicalism is a metaphysical thesis asserting that the universe has some kind of unified dynamic structure. Explicating unity in that way invites the charge of circularity, a charge that is not actually valid (see Maxwell 1998, pp. 118-23 and 168-72). The account given in this chapter forestalls this charge from the outset.

composed of physical entities interacting by means of different forces. For (5), sub-regions contain systems composed of entities interacting by means of the same force, but with different dynamical properties such as values of mass or charge. For (6), sub-regions contain systems composed of different entities which cannot be transformed into each other by means of symmetry operations. For (7), sub-regions contain systems composed of different entities which cannot be construed to differ only because of the product of spontaneous symmetry breaking. For (8), S contains one sub-system consisting of empty spacetime, and another consisting spacetime plus some physical entity, and the one cannot be transformed into the other by means of a symmetry operation. We have here eight facets of a single conception of unity.²¹²

It needs now to be appreciated that, corresponding to these eight facets of unity, (1) to (8), there are eight different metaphysical theses, eight different versions of physicalism, any one of which might be held to be the best choice of presupposition for physics.

13. This point is of fundamental importance for the problem of induction. Traditionally, the problem is interpreted as the problem of justifying exclusion of empirically successful theories that are *ad hoc* in sense (1): How can evidence from the past provide grounds for any belief about the future? This makes the problem seem highly "philosophical", remote from any problem realistically encountered in scientific practice. But the moment it is appreciated that the problem of justifying exclusion of empirically successful theories that are *ad hoc* in sense (1) is just an extreme, special case of the more general problem of excluding empirically successful theories that are *ad hoc* in senses (1) to (8), it becomes clear that this latter problem is a scientific problem, a problem of theoretical physics itself. For the implications of this crucial insight, and for a proposal as to how the problem of induction is to be solved exploiting it, see (Maxwell 1998, especially chs. 4 and 5). See also chapter 12 of the present work, and (Maxwell, 2004, appendix, section 6; and 2007, ch. 14, section 6.)

If T is the true theory of everything,²¹³ then we have eight different theses, each of the form “T is unified up to sense (n)” where $n = 1, 2, \dots, 8$, corresponding to the eight different kinds of unity. Let us call these eight theses “physicalism(n)”, where $n = 1, 2, \dots, 8$. It is assumed, here, that in each case, (1) to (8), the *degree* of unity, N, is 1. If we allow $N = 1, 2, 3, \dots$, then there are not eight, but infinitely many different versions of physicalism, depending on the degree of unity, N, that is asserted for any value of $n = 1, \dots, 7$. ($n = 8$ is exceptional in this respect in that, in this case, N can only equal 1 or 2, depending on whether spacetime and matter are, or are not, unified.²¹⁴) The different versions of physicalism can be specified to be: physicalism(n, N), with $n = 1 \dots 7$ and $N = 1, 2 \dots \infty$, or with $n = 8$ and $N = 1, 2$. A two-dimensional grid is placed over an infinite set of metaphysical theses, distinct versions of physicalism corresponding to distinct appropriate values of the coordinates (n,N), these theses being ordered with respect to degrees of unity.²¹⁵

14. It may seem that there is rather a jump here, from T referring to any fundamental dynamical physical theory in (1) to (8), to T referring to a “theory of everything”. However, the proper way to apply (1) to (8) is to the totality of fundamental physical theory (whether this consists of many or just one theory), and thus, in a sense, to candidate “theories of everything” (If a range of phenomena has no theory, then empirical laws governing these phenomena must be treated as theories.). If we do not do this, disunity could always be evaded, as far as (4) to (7) are concerned at least, by chopping a theory disunified to degree N into N distinct unified theories. When it comes to non-empirical considerations governing choice of theory, what matters is the way individual theories fit into the totality of fundamental physical theory – the degree of unity of the whole of fundamental physical theory.

15. Although, even in this case, one could imagine that there are different degrees of unification of spacetime and matter.

16. Is physicalism(n,N) to be interpreted so as to be compatible with stronger versions of physicalism, such as physicalism(n+1, N-1) – assuming here that $n < 8$ and $N > 1$? If we want the different versions of physicalism to constitute metaphysical theses that are, as far as possible, mutually exclusive, then we should interpret physicalism(n,N) to be incompatible with stronger, more unified versions of physicalism. But if we want physicalism to play the role in physics of excluding more or less disunified theories, then it will be convenient to interpret physicalism(n,N) to be compatible with stronger, more unified versions of the doctrine. In this second case, physicalism(n,N) has the role of excluding theories more disunified than (n,N), but not theories more unified than this. In

It deserves to be noted in passing that there is a close connection between the “unity” of a physical theory and its “explanatory power”. Explanatory power, one might say, is unity plus empirical content. The explication of “unity” indicated here and spelled out in more detail in chapter nine is thus also an explication of “explanatory power”.²¹⁶

4 – Conflicting Desiderata for Acceptability of Metaphysical Theses

Physics must exclude empirically successful disunified theories from consideration if theoretical knowledge in physics is to be possible at all. In persistently excluding such disunified theories, physics thereby makes a persistent metaphysical presupposition, as we saw in section 2. But what ought this presupposition to be? Section 3 has revealed that there are at least eight candidates, namely physicalism(n) with $n = 1, \dots, 8$ (and, potentially, many more if $N > 1$ for some n .) Which of these is the best choice if physics is to comply with PIR?

This question is particularly hard to answer because conflicting desiderata arise when it comes to considering what metaphysical thesis physics should accept.

On the one hand it is reasonable to argue that that thesis should be accepted which can be shown to be the most conducive to progress in theoretical physics so far. In the next section I will demonstrate that this picks out the relatively specific and contentful thesis of physicalism(8) with $N = 1$.

On the other hand, however, it is reasonable to argue that that thesis should be accepted which has the least content which is just

what follows, physicalism is to be interpreted in this latter way, as we will be considering the role physicalism has in excluding disunified theories from physics.

17. This explication of the “unity” and “explanatory power” of theories improves on proposals put forward by Friedman (1974), Kitcher (1981), Watkins (1984) and others. For a critical assessment of these and other proposals see Salmon (1989). For my criticism of the proposals of Friedman, Kitcher and Watkins and others, see Maxwell (1998, pp. 61-8). This book also contains a detailed account of my positive theory of explanatory power: see especially chs. 3 and 4.

sufficient to exclude the empirically successful disunified theories that current methods of physics do exclude. We have almost no grounds for holding that any version of physicalism is true, or is more likely to be true than some other version. Whatever we choose, we are very likely to choose a thesis that is false. Our best bet, then, is to choose that thesis which has the least possible content which suffices to exclude those disunified theories that are excluded from physics, since the less the content of a thesis – other things being equal – the more likely it is to be true. As we shall see in section 6 below, this leads to a choice quite different from physicalism(8).

Which of these conflicting lines of argument should be accepted?

In answering this question, I proceed as follows. In section 5, I spell out the argument for holding that physicalism(8) is the most fruitful version of physicalism for physics. In section 6, I spell out the argument for accepting the minimal version of physicalism. And in section 7, I argue that a new conception of physics resolves the conflict.

5 – Empirically Fruitful Metaphysics

Before I plunge into my argument for accepting physicalism(8), there is a preliminary question I must answer: Why does the unique fruitfulness of physicalism(8) for physics, supposing it can be established, provide grounds for its acceptance?

The basic idea of PIR, as it applies to physics, is that substantial, problematic, influential and implicit metaphysical assumptions need to be made explicit so that they can be critically assessed, so that alternatives can be developed and considered, in the hope that assumptions more conducive to progress can be developed and accepted. In other words, according to PIR, that assumption ought to be accepted which seems to be the most conducive to progress in theoretical physics.

It is important to appreciate just how profoundly influential over the success or failure of theoretical physics choice of metaphysical thesis, of the kind we are considering, is likely to be. This influence is exercised in two ways. First, the metaphysical

presupposition of a community of physicists influences – even determines – the direction in which physicists look in order to develop new physical theories. If physicists are convinced that the universe is made up of point-atoms which interact by means of centrally-directed, rigid forces – as many were for much of the 19th century – then physicists will persistently seek to develop theories which postulate such entities. Physicists who believe that the basic stuff of the universe is energy may be prompted to develop theories of a rather different type. Secondly, and even more important, the metaphysical presupposition of physics, implicit in non-empirical methods of physics, influences – or co-determines (with evidence) – what theories are accepted and rejected. The success or failure of physics will be highly dependent on whether the non-empirical methods adopted – and thus the corresponding metaphysical theses presupposed – are, or are not, conducive to the selection of theories capable of meeting with empirical success. Adopting the methodological principle “accept only theories that postulate atoms” amounts to presupposing that the universe is made up of atoms (or at least behaves, to a high degree of approximation, as if it is): if this is correct, this presupposition and associated methodological principle may well lead to empirical success. But if the universe is not made up of atoms, and does not even behave as if it is, adopting this methodological principle and presupposing the associated metaphysical thesis is likely to stifle severely scientific progress.

In short, physics must make some metaphysical presupposition for there to be any theoretical knowledge in physics at all. Since the metaphysical theses in question are about the ultimate nature of the universe, the domain of our ignorance, whatever we assume is almost bound to be false. Accepting a false assumption is likely to severely stifle progress in the theoretical physics. It matters enormously, for the progress of physics, that a good choice of metaphysical thesis is made. Just about the only grounds we have for preferring one thesis to another is that one seems to be more conducive to progress in theoretical physics than another. Thus we ought to prefer that thesis which seems to be the most conducive to progress in physics. This is the choice physics needs to make in order to comply with the requirement of rigour of PIR.

Here, now, are the reasons for holding that physicalism(8) should be the preferred metaphysical thesis for physics, largely because this thesis has proved to be more fruitful for progress in physics than any rival thesis.

First, it deserves to be noted that what needs to be made explicit and accepted, if physics is to comply with PIR, is that thesis which is implicit in the current non-empirical *methods* of physics – methods which determine which theories are to be accepted and rejected on grounds of simplicity, unity, explanatoriness. There can be no doubt that, as far as non-empirical considerations are concerned, the more nearly a new fundamental physical theory satisfies all eight of the above requirements for unity, with $N = 1$, the more acceptable it will be deemed to be. Furthermore, failure of a theory to satisfy elements of these criteria is taken to be grounds for holding the theory to be false even in the absence of empirical difficulties. For example, high energy physics in the 1960s kept discovering more and more different hadrons, and was judged to be in a state of crisis as the number rose to over one hundred. Again, even though the standard model (the current quantum field theory of fundamental particles and forces) does not face serious empirical problems, it is nevertheless regarded by most physicists as unlikely to be correct just because of its serious lack of unity. In adopting such non-empirical criteria for acceptability, physicists thereby implicitly assume that the best conjecture as to where the truth lies is in the direction of physicalism(8). PIR requires that this implicit assumption – or conjecture – be made explicit so that it can be critically assessed and, we may hope, improved. Physics with physicalism(8) explicitly acknowledged as a part of conjectural knowledge is more rigorous than physics without this being acknowledged because physics pursued in the former way is able to subject non-empirical methods to critical appraisal as physicalism(8) is critically appraised, whereas physics pursued in the latter way cannot do this.

The really important point, however, in deciding what metaphysical assumption of unity to accept, is that what needs to be considered is not just current theoretical knowledge, or current methods, but the whole way theoretical physics has developed

during the last 400, or possibly 2,000 years. The crucial question is this: What metaphysical thesis does the best justice to the way theoretical physics has developed during this period in the sense that successive theories increasingly successfully exemplify and give precision to this metaphysical thesis in a way which no rival thesis does? The answer is physicalism(8), as the following considerations indicate.

All advances in theory in physics since the scientific revolution have been advances in unification, in the sense of (1) to (8) above. Thus Newtonian theory (NT) unifies Galileo's laws of terrestrial motion and Kepler's laws of planetary motion (and much else besides): this is unification in senses (1) to (3). Maxwellian classical electrodynamics, (CEM), unifies electricity, magnetism and light (plus radio, infra red, ultra violet, X and gamma rays): this is unification in sense (4). Special relativity (SR) brings greater unity to CEM, in revealing that the way one divides up the electromagnetic field into the electric and magnetic fields depends on one's reference frame: this is unification in sense (6). SR is also a step towards unifying NT and CEM in that it transforms space and time so as to make CEM satisfy a basic principle fundamental to NT, namely the (restricted) principle of relativity. SR also brings about a unification of matter and energy, via the most famous equation of modern physics, $E = mc^2$, and partially unifies space and time into Minkowskian spacetime. General relativity (GR) unifies spacetime and gravitation, in that, according to GR, gravitation is no more than an effect of the curvature of spacetime – a step towards unification in sense (8). Quantum theory (QM) and atomic theory unify a mass of phenomena having to do with the structure and properties of matter, and the way matter interacts with light: this is unification in senses (4) and (5). Quantum electrodynamics unifies QM, CEM and SR. Quantum electroweak theory unifies (partially) electromagnetism and the weak force: this is (partial) unification in sense (7). Quantum chromodynamics brings unity to hadron physics (via quarks) and brings unity to the eight kinds of gluons of the strong force: this is unification in sense (6). The standard model unifies to a considerable extent all known phenomena associated with fundamental particles and the forces between them (apart from gravitation): partial unification in senses

(4) to (7). The theory unifies to some extent its two component quantum field theories in that both are locally gauge invariant (the symmetry group being $U(1) \times SU(2) \times SU(3)$). All the current programmes to unify the standard model and GR known to me, including string theory or M-theory, seek to unify in senses (4) to (8).²¹⁷

In short, all advances in fundamental theory since Galileo have invariably brought greater unity to theoretical physics in one or other, or all, of senses (1) to (8): all successive theories have increasingly successfully exemplified and given precision to physicalism(8) to an extent which cannot be said of any rival metaphysical thesis, at that level of generality. The whole way theoretical physics has developed points towards physicalism(8), in other words, as the goal towards which physics has developed. Furthermore, what it means to say this is given precision by the account of theoretical unity given in section 3 above.

In assessing the relative fruitfulness of two rival metaphysical theses, M_a and M_b , for some phase in the development of theoretical physics that involves the successive acceptance of theories $T_1, T_2, \dots T_n$, two considerations need to be born in mind. First, how potentially fruitful are M_a and M_b , how specific or precise, and thus how specific in the guidelines offered for the development of new theories? Second, how actually fruitful are M_a and M_b , in the sense of how successful or unsuccessful has the succession of theories, $T_1, T_2, \dots T_n$, been when regarded as a research programme with M_a or M_b as its key idea? When both considerations are taken into account, physicalism(8) comes out as more fruitful for theoretical physics from Newton to today than any rival thesis (at its level of generality). Physicalism(7) is not as specific as physicalism(8), and thus not as potentially fruitful, does not do justice to the way GR absorbs the force of gravitation into the nature of spacetime, and does not do justice to current research programmes which seek to unify matter and spacetime. (All of physicalism(n), $n = 1, 2, \dots 7$, are scientifically fruitful to some extent, but decreasingly so as n goes down from 7 to 6... to 1, in

18. For further discussion see (Maxwell 1998, pp. 80-89, 131-40, 257-65 and additional works referred to therein).

view of the decreasing specificity and content of these versions of physicalism.)

The notion of “research programme” appealed to here is similar to, but not the same as, the notion developed by Lakatos (1970). The main differences are as follows. For Lakatos, the “hard core” of a research programme was a testable theory rendered metaphysical by a methodological decision; the main research activity associated with a research programme involved developing successful applications of the theory, guided by the “positive heuristic” stemming from the “hard core”. (In all this, Lakatos followed Kuhn’s conception of “normal science”, giving Lakatosian terms to Kuhnian ideas: see Kuhn, 1970.) In the text, I have assumed that the metaphysics of a research programme is authentic, inherently untestable metaphysics, the main research task being to develop a succession of theories which progressively capture the metaphysics more and more successfully. The account of “degrees of disunity” given in section 3 above provides a precise way of assessing the extent to which successive “totalities of fundamental physical theory” do, or do not, increasingly successfully capture physicalism. Thus, given such a succession, $T_1, T_2, \dots T_m$, with degrees of disunity $N_1, N_2, \dots N_m$, of type (5-7), with $N_1 > N_2 > \dots > N_m$, then $T_1, T_2, \dots T_m$ do progressively capture physicalism(5-7) more and more successfully. There is nothing like this in the Lakatosian account of research programme, lacking as it does the solution to the problem of unity of theory. Finally, there is a substantial difference in the intended application of the two notions. Whereas I see science as a whole as one gigantic research programme, the hierarchy of versions of physicalism being presupposed as the metaphysical “hard core”, it is essential to Lakatos’s quasi-Popperian conception of science that science is made up of competing research programmes. This means that, for Lakatos, science cannot be viewed as one gigantic Lakatosian research programme (since, if it were, there could be no competitor). Lakatos does say, it is true “Even science as a whole can be regarded as a huge research programme with Popper’s supreme heuristic rule; ‘devise conjectures which have more empirical content than their predecessors’” (Lakatos, 1970, p. 132). But there is here no overall Lakatosian “hard core” or “positive

heuristic". This is Popper's conception of science and, for Lakatos, his own conception of research programme is strictly inapplicable to science as a whole. For a more detailed comparison and critical assessment of the two views, see chapter four, section 8).²¹⁸

Some philosophers of science hold that the successive revolutions in theoretical physics that have taken place since Galileo or Newton make it quite impossible to construe science as steadily and progressively honing in on some definite view of the natural world (Kuhn, 1970a; Laudan, 1980). If attention is restricted to standard empiricism and physical *theory*, this may be the case. But the moment some form of presuppositionism is accepted, and one considers metaphysical theses implicit in the methods of science, a very different conclusion emerges. *All theoretical revolutions since Galileo exemplify the one idea of unity in nature*. Far from obliterating the idea that there is a persistent thesis about the nature of the universe in physics, as Kuhn and Laudan suppose, all theoretical revolutions, without exception, do exactly the opposite in revealing that theoretical physics draws ever closer to capturing the idea that there is an underlying dynamic unity in nature, as specified by physicalism(8).

There is a further point to be made in favour of physicalism(8). So far, every theoretical advance in physics has revealed that theories accepted earlier are false. Thus Galileo's laws of terrestrial motion and Kepler's laws of planetary motion are contradicted by Newtonian theory, in turn contradicted by special relativity, in turn contradicted by general relativity. The whole of classical physics is contradicted by quantum theory, in turn contradicted by quantum field theory. Science advances from one false theory to another. Viewed from a standard empiricist perspective, this seems discouraging and has prompted the view that all future theories will be false as well, a view which has been called "the pessimistic induction" (Newton-Smith, 1981, p. 14). Viewed from the perspective of science presupposing physicalism(8), however, this mode of advance is wholly encouraging, since it is required if physicalism(8) is true. Granted physicalism(8), the only way a dynamical theory can be precisely true of any restricted range of phenomena is if it is such as to be straightforwardly generalizable so as to be true of all phenomena. Any physical theory inherently restricted to a limited range of phenomena, even though containing a wealth of true approximate predictions about these phenomena, must nevertheless be

19. Note 4 above indicates how my view differs from Popper's "metaphysical research programmes".

strictly false: only a theory of everything can be a candidate for truth!

Not only does the way physics has advanced from one false theory to the next accord with physicalism(8). The conception of unity sketched in section 3 successfully accounts for another feature of the way theoretical physics has advanced. Let T_1, T_2, \dots, T_n stand for successive stages in the totality of fundamental theory in physics. Each of T_1, T_2, \dots, T_n contradicts physicalism(8), in that each of T_1 etc. asserts that nature is *disunified*, whereas physicalism(8) asserts that it is *unified*. This might seem to make a nonsense of the idea that T_1, T_2, \dots, T_n is moving steadily and progressively towards some future T_{n+r} which is a precise, testable version of physicalism(8). But what section 3 shows is that, even though all of T_1, T_2, \dots, T_n are incompatible with physicalism(8), because disunified, nevertheless a precise meaning can be given to the assertion that T_{r+1} is closer to physicalism, or more unified, than T_r . This is the case if T_{r+1} is (a) of greater empirical content than T_r (since these are candidate theories of everything), and (b) of a higher degree of unity than T_r in ways specified in section 3. Thus the account of unity given above, involving physicalism(1 to 8), gives precision to the idea that a succession of false theories, $T_1 \dots T_n$, all of which contradict physicalism(8), nevertheless can be construed as moving ever closer to the goal of specifying physicalism(8) as a precise, testable, physical theory of everything.

6 – Metaphysical Minimalism

I turn now to the argument designed to show that that version of physicalism should be accepted which is the weakest available which just suffices to exclude theories more disunified than currently accepted physical theories – at present the standard model (SM) and general relativity (GR). We may take this to be the strongest version of physicalism that is compatible with SM+GR. This, it may be argued, leads to physicalism($n = 3$) being accepted – or possibly physicalism($n = 4, N = 4$), where $N = 4$ does justice to the fact that SM+GR includes four distinct forces, electromagnetism, the weak and strong forces, and gravitation (electromagnetism and the weak force being only partially unified in quantum electroweak theory).

But physicalism($n = 3$) is, it may be argued, much too strong. It assumes that SM+GR is consistent. Quantum field theory ordinarily assumes flat Minkowskian spacetime, but it can be extended to apply in the curved spacetime of GR. Problems arise, however, when attempts are made to treat quantum mechanically described matter or energy as the source of gravitation, and so of the curvature of spacetime, according to GR. No version of

SM+GR can be applied satisfactorily to the interior of black holes. Insurmountable problems arise if one tries to incorporate the quantum measurement of orthodox quantum theory into the framework of GR. SM+GR is, it seems, not a consistent theory. Even GR on its own faces a problem of consistency, in that GR predicts that a singularity forms inside a black hole, which constitutes a breakdown of the continuity of spacetime, and thus a kind of inconsistency of the theory. In order to assess the degree of disunity of SM+GR one has to consider a version patched up in an ad hoc way so as to create a theory that is at least *consistent*. Such an ad hoc patching up will further increase the kind and degree of disunity of the theory.

There is a further point. Orthodox quantum theory (OQT) holds that electrons, atoms, even molecules can be in superpositions of states – at different spatial locations at the same time, for example – whereas macroscopic measuring instruments cannot (measuring instruments detecting *one* outcome and not a *superposition* of outcomes). This means OQT holds that different laws hold at different levels – different masses, levels of complexity, or whatever it may be. This in turn means that OQT is disunified in way $n = 2$. Physicalism($n = 2, N = 2$) is perhaps compatible with OQT, but any more unified version of physicalism is not. This is a very serious level of disunity. What makes the matter even worse is that OQT does not specify what variables are involved – let alone what value of what variables – in the transition from quantum states and superpositions to classical states without such superpositions.²¹⁹

7 – The Hierarchical View

Should physics accept physicalism(8) in line with the argument of section 5, and risk committing physics to a highly specific and

20. See chapter 13 for further discussion of these issues.

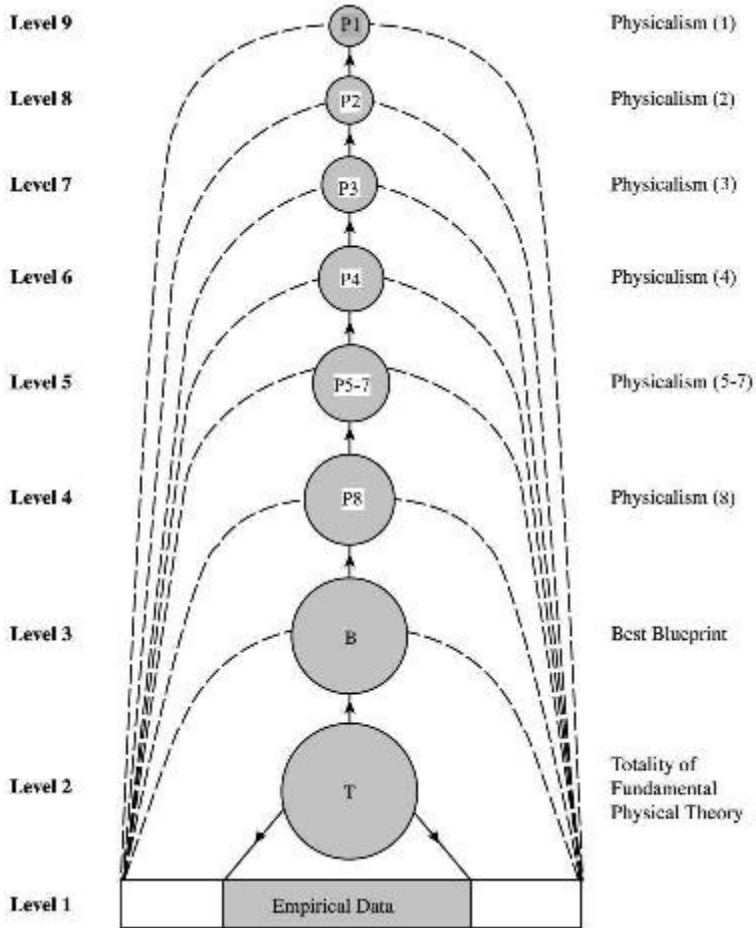


Diagram 3: Another version of Aim-Oriented Empiricism (AOE)

contentful version of physicalism all too likely to be false, despite its fruitfulness for physics up to the present? Or should physics accept some clumsy version of physicalism indicated in section 6, less contentful and thus more likely to be true, but entirely lacking in fruitful guidelines for the development of new physical theories?

No version of presuppositionism which restricts itself to adopting a single (if composite) metaphysical thesis can resolve the conflicting desiderata satisfactorily, that are highlighted in

these two questions. But this conflict is resolved if we adopt a version of presuppositionism which holds that we need to see physics as adopting a *hierarchy* of theses, from physicalism(8) near the bottom of the hierarchy to physicalism(1) at the top: see diagram 3. Physicalism(5-7) are on the same level since they are all but equivalent to one another. As we descend the hierarchy, from level (9) to level (3), theses become increasingly contentful and specific, increasingly potentially fruitful for future progress in theoretical physics but also increasingly likely to be false and in need of revision. As one moves from level (9) to level (3), the corresponding methodological requirements for unity, depicted as sloping dotted lines in diagram 3, become increasingly demanding, but also increasingly speculative and uncertain. The totality of physical theory, at any given stage in the development of physics (except when a candidate unified theory of everything has been proposed and accepted) will only satisfy these methodological rules partially; a new theory, in order to be an advance from the standpoint of unity, must lead to a new totality of theory satisfying the methodological rules better than the previous totality.

This hierarchical view has the following advantages over any version of presuppositionism which restricts itself to a single (possibly composite) thesis. First, the hierarchical view does justice to *both* apparently conflicting desiderata, indicated above, which cannot be done if a single metaphysical assumption is made. The hierarchical view includes *both* the uniquely scientifically fruitful thesis of physicalism(8) and the much less specific and problematic theses of physicalism (1) or (2). Second, the hierarchical view, as a result of making explicit metaphysical theses implicitly presupposed in adopting methods associated with levels (4) to (8), facilitates criticism and revision of these methods, which may well need to be done at some stage (if the corresponding metaphysical theses are false). Such criticism and revision is not facilitated if a single thesis is presupposed. Third, the hierarchical view assists revision of the more contentful and specific versions of physicalism low down in the hierarchy by providing a framework of relatively unproblematic assumptions and methods, at levels (9) to (6), which place restrictions on the way the more specific, problematic versions of physicalism may be

revised, should the need to do so arise. If a succession of increasingly empirically successful theories are developed, T_1, T_2, \dots , all of which clash with physicalism(8), but which accord increasingly well with physicalism (7), this might be taken as grounds for rejecting or modifying physicalism(8).

The reasons given above for including the relatively specific, scientifically fruitful metaphysical thesis of physicalism(8) in the hierarchy of accepted theses are reasons also for accepting an even more specific, scientifically fruitful metaphysical thesis, should one be available. A glance at the history of physics reveals that a succession of much more specific metaphysical theses have been accepted, or taken very seriously, for a time, each thesis being an attempt to capture aspects of physicalism. Ideas at this level include: the universe is made up of rigid corpuscles that interact by contact; it is made up of point-atoms that interact at a distance by means of rigid, spherically symmetrical forces; it is made up of a unified field; it is made up of a unified quantum field; it is made up of quantum strings. These ideas tend to reflect the character of either the current best accepted physical theory, or assumptions made by current efforts to develop a new theory. This is not sufficient to be scientifically fruitful in the way that physicalism (8) is. For this, we require that the thesis in question is such that all accepted fundamental physical theories since Newton can be regarded as moving steadily towards capturing the thesis as a testable physical theory, in the manner indicated in section 3. One candidate for such a thesis is the following:

Lagrangianism: the universe is such that all phenomena evolve in accordance with Hamilton's principle of least action, formulated in terms of some unified Lagrangian (or Lagrangian density), L . We require, here, that L is not the sum of two or more distinct Lagrangians, with distinct physical interpretations and symmetries, for example one for the electroweak force, one for the strong force, and one for gravitation, as at present; L must have a single physical interpretation, and its symmetries must have an appropriate group structure (the group not being a product of sub-groups). We

require, in addition, that current quantum field theories and general relativity emerge when appropriate limits are taken.²²⁰

All accepted fundamental physical theories, from Newton on, can be given a Lagrangian formulation. Furthermore, if we consider the totality of fundamental physical theory since Newton (empirical laws being included if no theory has been developed) then, as in the case of physicalism(8), every new accepted theory has brought the totality of physical theory nearer to capturing Lagrangianism. Thus Lagrangianism is at least as scientifically fruitful as physicalism(8). In fact it is *more* scientifically fruitful since it is very much more specific and contentful. The reasons for accepting physicalism(8) are reasons for accepting Lagrangianism too as the lowest thesis in the hierarchy of metaphysical theses, very much more potentially scientifically fruitful than physicalism(8), but also more speculative, more likely to need revision: see diagram 3.

It deserves to be noted that something like the hierarchy of metaphysical theses, constraining acceptance of physical theory from above, is to be found at the empirical level, constraining acceptance of theory from below. There are, at the lowest level, the results of experiments performed at specific times and places. Then, above these, there are low-level experimental laws, asserting that each experimental result is a repeatable effect. Next up, there are empirical laws such as Hooke's law, Ohm's law or the gas laws. Above these there are such physical laws as those of electrostatics or of thermodynamics. And above these there are theories which have been refuted, but which can be "derived", when appropriate limits are taken, from accepted fundamental theory – as Newtonian theory can be "derived" from general relativity. This empirical hierarchy, somewhat informal perhaps, exists in part for precisely the same epistemological and methodological reasons I have given for the hierarchical ordering of metaphysical theses: so that relatively contentless and secure theses (at the bottom of the hierarchy) may be distinguished from more contentful and insecure theses (further up the hierarchy) to

21. Lagrangianism is discussed in Maxwell (1998, pp. 88-9).

facilitate pinpointing what needs to be revised, and how, should the need for revision arise. That such a hierarchy exists at the empirical level provides further support for my claim that we need to adopt such a hierarchy at the metaphysical level.

8 – Aim-Oriented Empiricism

The hierarchical view depicted in diagram 3 may need to be rejected in its entirety as physics advances. If we exclude from consideration physicalism($n = 1, N = \infty$) which permits anything, the hierarchical view assumes that the universe is at least partially physically comprehensible in the sense that phenomena occur in accordance with physical laws which are more or less disunified, the traditional distinction between laws and initial conditions being presupposed. But even though the universe is physically comprehensible, the traditional distinction between laws and initial conditions might not be observed. The true theory of everything might be cosmological in character, and might specify unique initial conditions for the universe.²²¹ This possibility, and other possibilities of this kind, could no doubt be accommodated within a modified version of the above hierarchical view. But there are other possibilities, of philosophical interest even if of no interest to physics as at present constituted, which cannot be so accommodated. Perhaps God is ultimately responsible for all natural phenomena, or some kind of cosmic purpose or cosmic programme analogous to a computer programme (as has been suggested). In these cases the universe would be comprehensible but not physically comprehensible – even though it might mimic a physically comprehensible universe, to some extent.

In order to accommodate these, and other such, possibilities we need to embed the above hierarchical view in the broader view I have called “aim-oriented empiricism” (AOE), depicted in diagram 2. However, in order to do this, AOE, as depicted in diagram 2, needs to be modified. In this modified diagram, we would have, in succession, as we go up the hierarchy: empirical phenomena (level 1); accepted fundamental physical theory (level 2); best blueprint (level 3); then, at level 4, the thesis that the universe is perfectly

22. See my (2004, pp. 198-205) for a suggestion along these lines.

comprehensible physically – physicalism($n=8$); at level 5 the thesis that the universe is all-but perfectly comprehensible physically – physicalism($n>4, N=1$); at level 6, the thesis that the universe is partially comprehensible physically – physicalism($n>4, N<10$); the thesis that the universe is partially comprehensible, either physically or in some other way (level 7); the thesis of meta-knowability: the universe is such that we can discover how to improve our methods of learning (level 8); and, finally, the thesis of partial knowability: the universe is such that we can acquire some knowledge of our local circumstances (level 9).

The level 8 thesis, here, asserts that the universe is such that there is some rationally discoverable proposition about its nature (relative to existing knowledge) which, if accepted, makes it possible progressively to improve methods for the improvement of knowledge. "Rationally discoverable" means at least that the thesis is not an arbitrary choice from infinitely many analogous theses. This thesis is to be interpreted as asserting that the universe is not epistemologically malicious, in the sense that apparently improved methods lead to apparent new knowledge which turns out, subsequently, to be illusory, there being no possibility of discovering this before it is revealed. (This is clarified in the next section.) Level 9 asserts that the universe is such that we can continue to acquire knowledge of our local circumstances, sufficient to make life possible.

Such an amalgam of diagrams 2 and 3 is not altogether satisfactory. It is somewhat arbitrary to declare that physicalism($n>4, N=9$) represents a partially physically comprehensible universe, but physicalism($n>4, N=10$) does not. Some may question that the alternative thesis at level 7 – the thesis that the universe is comprehensible non-physically – has anything to do with physics, or science. One can imagine circumstances, however, in which this thesis might well be preferred to any thesis of partial physical comprehensibility. Suppose God manifests Himself in the sky, and responds to our requests to perform specific miracles by doing just that – converting the orbit of Mercury into a square orbit, for example, or lifting everyone on earth one foot into the air, except where this is impossible or

dangerous. How could we not accept, in such circumstances, that God is ultimately in charge of Nature?

Reasons for accepting these theses are similar to those given for accepting the hierarchy of theses of the view depicted in diagrams 2 and 3. As a result of accepting this hierarchy of theses, physics provides itself with a framework of assumptions, and associated methods, high up in the hierarchy, which it will never benefit the pursuit of knowledge to reject; in this way, a framework is created within which very much more substantial and dubious theses, low down in the hierarchy, can be critically scrutinized and, we may hope, improved. The thesis at the top of the hierarchy, at level 9, asserts that the universe is such that we can acquire some knowledge of our local circumstances. If this assumption is false, we will not be able to acquire knowledge whatever we assume. We are justified in accepting this assumption permanently as a part of our knowledge, even if we have no grounds for holding it to be true, since accepting it can only help, and cannot hinder, the acquisition of knowledge whatever the universe is like. The thesis at level 4, physicalism(8), deserves to be accepted because of its extraordinary scientific fruitfulness. All major theoretical developments in physics point towards, and draw closer to, physicalism(8), in that they are invariably major steps in unification, as we saw in section 5 above. From the standpoint of scientific fruitfulness, at this level, physicalism(8) has no rival, unless it is the even more substantial thesis of Lagrangianism. Nevertheless, physicalism(8) may be false, and physics may, at some stage, need to adopt a different thesis. Accepting physicalism(8) within the framework of theses just indicated facilitates the development of alternatives, should this be necessary.

The view depicted in diagram 3 has, perhaps, a more direct relevance to theoretical physics; AOE would become relevant if it emerged that the universe differs radically from the way modern science assumes it to be. AOE is more relevant to the philosophy of physics; it is required to solve the problem of induction, as we shall see in chapter 12. (See also Maxwell, 1998, ch. 5; 2007, ch. 14.) It is also required to rebut objections of circularity, as we shall now see.

9 - Circularity Problem Solved

One feature of the views depicted in diagrams 2 to 4 may be deemed to be puzzling. They hold that when metaphysical thesis and physical theory clash, physical theory may be revised, but also that metaphysical thesis may be revised. How is such a two-way influence possible? In what follows I consider AOE as depicted in the previous section.

The first point to note is that just such a two-way influence occurs when theory and experiment clash. In general, if a theory clashes with an experiment that has been subjected to expert critical scrutiny and repeated, the theory is rejected. But on occasions it turns out that it is the experimental result that is wrong, not the theory.²²² In a somewhat similar way, if a new theory increases the conflict²²³ between the totality of physical theory and the currently accepted metaphysical thesis, at level 3 of the previous section, the new theory will be rejected (or not even considered or formulated). On occasions, however, a new theory may be developed which increases the conflict between the totality of theory and the current thesis at level 3 but decreases the conflict between the totality of theory and physicalism at level 4. In this case the new theory may legitimately be accepted and the thesis at level 3 may be revised. In principle, as I have already indicated, theses even higher up in the hierarchy may legitimately be revised in this way. A virtue of these hierarchical views is that they make possible and facilitate such two-way revisions.

However, another, potentially more serious problem faces the two hierarchical views indicated above. Both incorporate what seems to be vicious circularity. Acceptance of theories is

23. In practice, physical theory is persistently used to correct clashing experimental results, in that theory is used to reveal that experimental equipment is not working properly.

24. What does “increase the conflict” mean here? It means that the kind or degree of unity of the totality of fundamental physical theory gets worse with respect to the currently accepted metaphysical thesis at level 3. (This thesis asserts that there is a certain *kind* of unity in nature. The more the totality of physical theory departs from this kind of unity, the greater the “conflict” with the thesis at level 3.)

influenced by their degree of accord with metaphysical principles, the acceptance of which is in turn, in part, influenced by an appeal to the empirical success of physical theories. The claim is that as theoretical knowledge and understanding improves, metaphysical theses and associated methods improve as well. There is something like positive feedback between improving knowledge, and improving knowledge-about-how-to-improve knowledge. This, it is claimed, is the methodological key to the great success of modern science, namely that it adapts its metaphysical assumptions and methods (its aims and methods) to what it finds out about the nature of the universe.²²⁴ But how can such a circular procedure conceivably be valid?

This is not an objection to the arguments and views put forward so far. No attempt has been made to justify claims to theoretical knowledge. The argument has been modest: granted acceptance of current physical theories and adoption of current methods, then physics is more rigorous (in that it accords better with PIR) if implicit metaphysical assumptions are made explicit, and those assumptions chosen which seem best to promote what we take to be scientific progress. The circularity objection would arise, however, if we were to go beyond the modest aspirations of sections 2 to 8, and attempt to solve the problem of induction,²²⁵ and justify acceptance of empirically successful unifying theories, within the context of AOE. Or, putting the matter slightly differently, AOE cannot be acceptable because the problem of induction cannot conceivably be solved within its framework: the moment the attempt is made to *justify* acceptance of scientific

25. See Maxwell (1974, especially part II) and, more recently, Maxwell (1998, 17-19). Others too have argued that the methods of science improve with improving knowledge, but have done so only within the framework of standard empiricism: see, for example, Boyd (1980).

26. No attempt is being made here, I hope it is clear, to solve the problem of induction. I merely seek to rebut the objection that the problem cannot conceivably be solved, granted AOE. In chapter 12, I will however argue that AOE solves the problem of induction – as I have argued elsewhere: see Maxwell (1998, ch. 5; 2004, appendix, section 6; and especially 2007, ch. 14, pp. 400-430).

theories and metaphysical theses as claims to knowledge, vicious circularity sets in. How is this circularity objection to be met?

Here, in a nutshell, is the answer. Permitting metaphysical assumptions to influence what theories are accepted, and at the same time permitting theories to influence what metaphysical assumptions are accepted, may (if carried out properly), *in certain sorts of universe*, lead to genuine progress in knowledge. The level 6 thesis of meta-knowability, of AOE, asserts that *this is just such a universe*. And furthermore, crucially, reasons for accepting meta-knowability make no appeal to the success of science. In this way, meta-knowability legitimises the potentially invalid circularity of AOE, and of the component view depicted in diagram 3.

Relative to an existing body of knowledge and methods for the acquisition of new knowledge, possible universes can be divided up, roughly, into three categories: (i) those which are such that the meta-methodology of AOE can meet with no success, not even apparent success, in the sense that new metaphysical ideas and associated methods for the improvement of knowledge cannot be put into practice so that success (or at least apparent success) is achieved; (ii) those which are such that AOE can meet with genuine success; and (iii) those which are such that AOE appears to be successful for a time, but this success is illusory, this being impossible to discover during the period of illusory success. Meta-knowability asserts that this universe is a type (ii) universe; it rules out universes of type (i) and (iii).

Meta-knowability asserts, in short, that the universe is such that AOE can meet with success and will not lead us astray in a way in which we cannot hope to discover by normal methods of scientific inquiry (as would be the case in a type (iii) universe). If we have good grounds for accepting meta-knowability as a part of scientific knowledge – grounds which do not appeal to the success of science – then we have good grounds for adopting and implementing AOE (from levels 8 to 3).

But what grounds are there for accepting the thesis of meta-knowability at level (6)? There are two.²²⁶

(a) Granted that there is *some* kind of general feature of the universe which makes it possible to acquire knowledge of our local environment (as guaranteed by the thesis at level 8), it is reasonable to suppose that we do not know all that there is to be known about what the *nature* of this general feature is. It is reasonable to suppose, in other words, that we can improve our knowledge about the nature of this general feature, thus improving methods for the improvement of knowledge. Not to suppose this is to assume, arrogantly, that we already know all that there is to be known about how to acquire new knowledge. Granted that learning is possible (as guaranteed by the level 8 thesis), it is reasonable to suppose that, as we learn more about the world, we will learn more about how to learn. Granted the level 8 thesis, in other words, meta-knowability is a reasonable conjecture.

(b) Meta-knowability is too good a possibility, from the standpoint of the growth of knowledge, not to be accepted initially, the idea only being reluctantly abandoned if all attempts at improving methods for the improvement of knowledge fail.

(a) and (b) are not, perhaps, very strong grounds for accepting meta-knowability; both are open to criticism. But the crucial point, for the present argument, is that these grounds for accepting meta-knowability, (a) and (b), are independent of the success of science. This suffices to avoid circularity.

If AOE lacks meta-knowability, its circular procedure, interpreted as one designed to procure justified knowledge, becomes dramatically invalid, as the following consideration reveals. Corresponding to the succession of accepted fundamental physical theories developed from Newton down to today, there is a succession of aberrant rivals which postulate that gravitation becomes a repulsive force from the beginning of 2050, let us say. Corresponding to these aberrant theories there is a hierarchy of aberrant versions of physicalism, all of which assert that there is an abrupt change in the laws of nature at 2050. The aberrant theories,

27. I here echo the two reasons for accepting meta-knowability given in chapter 4.

just as empirically successful as the theories we accept, render the aberrant versions of physicalism just as scientifically fruitful as non-aberrant versions of physicalism are rendered by the non-aberrant theories we actually accept. If we take it as given that we accept non-aberrant theories, the question of what reasons there are for rejecting empirically successful aberrant theories and associated aberrant versions of physicalism does not arise. But the moment we seek to *justify acceptance* of non-aberrant theories and *rejection* of aberrant theories, within the framework of AOE, the question of what reasons there are for rejecting aberrant theories and associated aberrant versions of physicalism arises. If AOE is bereft of meta-knowability, it is not easy to see what these reasons can be. But AOE with meta-knowability included does provide a reason: the aberrant versions of physicalism assert that this is a type (iii) universe, which violates meta-knowability.

Versions of physicalism(n) for which $n = 1$ or 3 and $N > 1$ would seem to violate meta-knowability. But other versions of physicalism with $N > 1$ need not clash with meta-knowability.²²⁷

10 – Conclusions

I have argued that if science is to be rigorous it needs to accept explicitly, as a core component of theoretical scientific knowledge, a hierarchy of metaphysical theses (and associated methods) concerning the dynamic unity, the physical comprehensibility, of the universe. I have shown that the notorious problem concerning the unity, the explanatory character, of physical theory can be solved within this hierarchical view of science. This solution, in turn, provides a precise way of assessing the scientific fruitfulness of rival metaphysical theses, from the standpoint of the empirical progressiveness of the research programmes to which they give rise.

28. Consider a universe such that progress in theoretical physics requires infinitely many theoretical revolutions, each revolution leading to the degree of unity of the totality of physical theory going up by one. In such a universe, meta-knowability is true, since it possesses a general feature which, once discovered, would aid progress in physics; nevertheless, it is also the case that physicalism($N = \infty$) is true. Physicalism holds, meta-knowability holds, but the universe is infinitely disunified, infinitely incomprehensible.

These results have dramatic implications for science, for our understanding of science, and for the relationship between science and philosophy. There is a major increase in the (acknowledged) *scope* of scientific knowledge. Whereas standard empiricism implies that science at present provides us with no knowledge about the ultimate nature of the universe (all current fundamental physical theories being false), the hierarchical view holds the opposite. Current science does include knowledge about the ultimate nature of the universe – knowledge which, though theoretical and conjectural, is nevertheless more secure than any accepted theory such as quantum theory or general relativity: physicalism(5-7) is true, even perhaps physicalism(8). Science becomes much more like natural philosophy, in that it incorporates sustained exploration and assessment of metaphysical theses, and associated methodological rules, as an integral, vital part of scientific research. Instead of metaphysics and philosophy being banished from science, they become a vital part of science.

Furthermore, as I have argued in detail elsewhere (Maxwell, 1976a, 1984a or 2007, 2000a, 2001a, 2004, 2010a), the arguments of this paper, when extended to take into account, not just the implicit metaphysical assumptions of science, but its implicit value and political assumptions as well, have even more dramatic implications not just for physics or natural science, but for social science, for the humanities, for academic inquiry as a whole.

Perhaps the time has come for philosophers to take note of these arguments which have such revolutionary implications for our intellectual landscape.

Chapter Eleven

Comprehensibility rather than Beauty

(Modified version of a paper posted on the Pittsburgh PhilSci Archive in 2004: [http://philsci-archive.pitt.edu/archive/00001770/.](http://philsci-archive.pitt.edu/archive/00001770/))

Abstract

Most scientists and philosophers of science recognize that, when it comes to accepting and rejecting theories in science, considerations that have to do with simplicity, unity, symmetry, elegance, beauty or explanatory power have an important role to play, in addition to empirical considerations. Until recently, however, no one has been able to give a satisfactory account of what simplicity (etc.) *is*, or how giving preference to simple theories is to be *justified*. But in the last few years, two different but related accounts have appeared, both of which address the above issues. On the one hand, James McAllister has argued that aesthetic criteria in science reflect scientists' judgements about what kind of theory is most likely to be empirically successful, based on the relative empirical success and failure of different kinds of theories in the past. Scientists employ what McAllister dubs "the aesthetic induction". On the other hand, I have argued that we need to see science as making a hierarchy of metaphysical assumptions about the comprehensibility and knowability of the universe, these assumptions asserting less and less as one ascends the hierarchy. One of the more substantial of these assumptions is that the universe is physically comprehensible. The key non-empirical feature a body of fundamental theories in physics must possess to be acceptable is unity. The better such a body of theory exemplifies the metaphysical thesis that the universe is physically

comprehensible, in the sense that it has a unified dynamic structure, so the more acceptable such a body of theory is, from this standpoint. This affects not just theoretical physics, but the whole of natural science. In this chapter I compare and contrast, and try to assess impartially the relative merits of, these two views.

1 Beauty or Comprehensibility?

Many scientists, and some philosophers of science, have acknowledged that aesthetic considerations play, quite properly, an important role in influencing acceptance and rejection of theories in science, in addition to empirical considerations. A famous example is Dirac, who went as far as to declare "It is more important to have beauty in one's equations than to have them fit experiment" (quoted in McAllister, 1996, p. 15).

The view that beauty ought to influence choice of theory in science faces, however, a serious problem. Why should beauty be a good indication of truth? Unless the truth is beautiful, and unless we have valid grounds for holding this to be the case, there can be no good reasons, it would seem, for giving preference to beautiful theories in science.

Not only may it seem dubious that we can have grounds for holding that the truth is beautiful; there may well seem to be grounds for holding that it is wildly implausible that the truth should be beautiful, especially in theoretical physics.

Whether we find something beautiful or ugly must depend, to some extent at least, on our personal, subjective, emotional responses to that thing. Aesthetic criteria have their roots deep in the human psyche, and in human culture. But physical reality, that which theoretical physics seeks to grasp, is utterly remote from the human psyche, from human culture. It may well seem utterly implausible that something as anthropomorphic, as personal, as quintessentially human and subjective, as ideas about beauty, should have anything to do with the ultimate nature of the physical universe, utterly impersonal and remote from the circumstances of human life. Beauty may seem to be the last consideration to take into account in assessing the merits of rival fundamental theories in physics.

An extremely interesting and original defence of the thesis that aesthetic considerations do quite properly influence theory choice in science has, however, been put forward recently by James McAllister (1996): see also his (1989), (1990) and (1991). Quite independently, I have, over a number of years, developed a view which resembles McAllister's view in a number of striking ways, but which is also different in important respects: see chapters 4 and

5, and especially chapters 9 and 10.²²⁸ In this chapter I compare and contrast the two views. I begin with a sketch, first of my view, then of McAllister's. I then discuss how they resemble, and differ from, each other. And finally I discuss the question of which is to be preferred.

2 Aim-Oriented Empiricism

The view I wish to defend, aim-oriented empiricism (AOE), has already been expounded in some detail in earlier chapters. Here, I will be very brief. AOE holds that science is obliged to make a big, persistent, metaphysical assumption about the nature of the universe. This assumption is implicit in those methods of science which specify that theories, in order to be accepted, must be sufficiently non-ad hoc, simple, unified or explanatory.

This claim is denied by the widely held view standard empiricism (SE), which asserts that, in science, all claims to knowledge are to be assessed impartially with respect to the evidence, *no thesis about the world being upheld permanently as a part of knowledge independently of evidence, let alone in violation of evidence*. Most, if not all, versions of SE stress that questions of simplicity, unity, beauty or explanatory power play a valid, important role in influencing choice of theory in science, in addition to considerations of empirical success – although some versions of SE give to simplicity, beauty or explanatory power much more important roles in science than other versions do. The decisive point that all versions of SE agree on is that no substantial thesis about the nature of the universe can be upheld as a part of scientific knowledge independently of empirical considerations, and certainly not in violation of empirical considerations. In so far as theory choice is biased in the direction of simplicity, unity, beauty or explanatoriness, this bias must not commit science to making the permanent assumption that nature herself is simple, unified, beautiful or explainable.

²²⁸. See also my (1972a, 1974, 1984a, 1993a, 2001a chapter 3 and appendix 3, 2004, 2007 chapter 14, and especially 1998; for recent summaries see my 1999a and 2000b).

This thesis of SE is common ground for logical positivism, inductivism, logical empiricism, hypothetico-deductivism, conventionalism, constructive empiricism, pragmatism, realism, induction-to-the-best-explanationism, and the views of Popper and Kuhn.²²⁹ McAllister too, as we shall see, defends a version of SE.

SE is, nevertheless, untenable, as we have already seen. Theoretical physics only ever accepts fundamental theories that are *unified* even though endlessly many empirically more successful disunified rivals can always be concocted. This means physics makes a big, permanent metaphysical assumption about the nature of the universe: it is such that, at the very least, all disunified theories are false.

At once, two questions arise: What precisely ought this assumption to be? How is it to be justified?

The solution to these two fundamental epistemological problem of science (the very existence of which is denied by SE) is to construe science as adopting, as a part of scientific knowledge, a hierarchy of assumptions about the comprehensibility and knowability of the universe, these assumptions asserting less and less about the universe as one ascends the hierarchy, thus being more and more likely to be true: see diagram 2 on page X. Corresponding to these cosmological assumptions there are methodological rules (represented by dotted lines in the diagram) which govern acceptance of assumptions lower down in the hierarchy, and which, together with empirical considerations, govern acceptance and rejection of scientific theories.

It is this conception of science, postulating more or less specific evolving aims and methods for science within a framework of more general fixed aims and methods, that I call *aim-oriented empiricism* (AOE).²³⁰ It is a special case of a more general idea of

²²⁹. For discussion of the claim that SE is upheld by a wide range of views about science see my (1998, pp. 37-45). Bayesianism might seem to reject SE, in acknowledging both prior and posteriori probabilities. But Bayesianism tries to conform to the spirit of SE as much as possible, by regarding prior probabilities as personal, subjective and non-rational, their role in theory choice being reduced as rapidly as possible by empirical testing: see Maxwell (1998, p. 44).

²³⁰. Corresponding to each cosmological thesis, at level 3 to 7, there is a more or less problematic *aim* for theoretical physics: to specify that cosmological thesis

aim-oriented rationality, according to which, whenever basic aims are problematic (as they usually are in science and in life) we need to display aims at distinct levels of specificity and generality, thus creating a framework of unproblematic, fixed aims-and-methods within which we have the best chance of improving more or less specific, problematic aims-and-methods as we proceed, in the light of success and failure.²³¹

For details concerning the theses at levels 1 to 7 of diagram 2 of page X, see chapters 4 and 5.

It may be objected that AOE is exclusively about theoretical physics, and cannot do justice to the variety of methods to be found in different branches of the natural sciences. In fact just the opposite is the case; as we saw in chapter 4, AOE predicts diversity of method throughout natural science, overlaid by unity of method at a meta-methodological level. AOE can do justice to the diversity of methods to be found in diverse sciences, without underlying unity and rationality being sacrificed.

Elsewhere I have argued at some length that AOE solves a number of outstanding problems about the nature of science, such as the problem of induction, the problem of what it means to say of a physical theory that it is unified, the problem of specifying precisely the nature of scientific method (just touched upon), the problem of verisimilitude, and the problem of how new fundamental physical theories can be discovered (see Maxwell, 1998, chapters 4-6; 2004, appendix; 2007, ch. 14); see too previous chapters of this book.

as a true, precise, testable, experimentally confirmed "theory of everything". The aim corresponding to level 7 is unproblematic: circumstances will never arise such that it would serve the interests of acquiring knowledge to revise this aim. As one descends the hierarchy of cosmological assumptions, the corresponding aims become increasingly problematic, increasingly likely to be unrealizable, just because the corresponding assumption becomes increasingly likely to be false. Whereas upper level aims and methods will not need revision, lower level aims and methods, especially those corresponding to level 3, will need to be revised as science advances. Thus lower level aims and methods evolve within the fixed framework of upper aims and methods.

²³¹. For the generalization of aim-oriented empiricism to form aim-oriented rationality see prologue and chapters 7 and 8. See also my (1984a or 2007; 2000a; 2001a, chapter 9; 2004, chapters 3 and 4; and 2010a, chapters 6 and 9).

How is acceptance of physicalism and the other metaphysical theses in the hierarchy of AOE to be justified? Here I can only sketch my answer; for a more detailed response see Maxwell (1998, especially ch. 5; 2007, ch. 14, section 6); and see next chapter.

The first point to note is that even our most trivial, everyday claims to factual knowledge contain implicit factual claims about the entire universe, and the ultimate nature of the universe. Thus, in claiming that I know I can walk across the room I am, implicitly, claiming to know that nowhere in the universe is there now occurring an explosion of chaos which will travel at almost infinite speed to engulf the earth, the room, and me before I can take a step. In claiming that I have the power to decide to walk across the room I am claiming, implicitly, that the ultimate nature of the universe is such that free will, in some meaningful sense, is both possible, and actual for me. Granted that even our most trivial common sense claims to knowledge contain cosmological and metaphysical dimensions, it should occasion no surprise that far more contentful scientific claims to knowledge do as well.

I have already argued that AOE, with its hierarchy of metaphysical theses at levels 3 to 7, is more rigorous than any version of SE, which would depict scientific knowledge as existing only at levels 1 and 2, everything above 2 being merely speculation to be considered in the context of discovery only. The AOE picture makes explicit and so criticizable and revisable assumptions that are substantial, influential and problematic, but only implicit granted SE. This in itself makes AOE more rigorous than SE, and provides a kind of justification for accepting metaphysical theses at levels 3, 4, and above.

SE tends to depict the transition from pre-science to science as the process of excluding metaphysical theses from science, only testable theories being candidates for scientific acceptance. But it is precisely this way of construing this transition that generates the unrigorous conception of science of SE, a symptom of this lack of rigour being the failure to solve the problem of induction – the problem of how theories can be confirmed by evidence. It is this, indeed, that *creates* the unsolvable problem of induction. What we need to do, instead, is to construe the transition from pre-science to

science as involving, ideally, not the ejection of untestable metaphysics from scientific knowledge altogether, but instead the articulation of implicit metaphysics as explicit theses, and then the selection of those metaphysical theses which either (a) the search for knowledge cannot do without, or (b) appear to be the most fruitful in leading to the growth of empirical knowledge (at levels 1 and 2). Above all, we need to organize these metaphysical theses into the hierarchical structure of AOE, thus creating a framework of relatively unchallengeable and unproblematic assumptions within which much more specific and problematic assumptions, lower down in the hierarchy, can be revised as empirical success and failure seem to require. In this way we can focus criticism on that part of the metaphysical presuppositions of science which, we conjecture, it is the most fruitful to criticize from the standpoint of achieving progress in scientific knowledge and understanding.

In other words, the fundamental epistemological problem of science is not the problem of induction, nor even the problem of justifying the truth of metaphysical assumptions made by science. It is, rather, the problem of showing that, from diverse metaphysical, cosmological assumptions that science might make, those that science actually selects are either indispensable for any attempt at acquiring knowledge, or are more fruitful for empirical progress than any rival assumptions.

The level 4 thesis of physicalism is to be accepted as a part of scientific knowledge, in short, because, at this level, and within the general framework of AOE, there is no other thesis that has proved as fruitful in promoting scientific progress at levels 1 and 2.

But is this correct? Is it not refuted by Kuhn's point (Kuhn, 1970a, chapter 13) that nothing theoretical survives a revolution, the new theory or paradigm being incommensurable with the old one?

This Kuhnian view, if valid at all, is most likely to be correct when applied to revolutions in fundamental theoretical physics, where radical discontinuity seems most marked. Ironically enough, it is above all here that Kuhn's claim fails. All revolutions in theoretical physics, despite their diversity in other respects, reveal one common theme: they are all gigantic steps in unification. Thus Newton unifies Kepler and Galileo. Maxwell's

theory of the electromagnetic field unifies electricity, magnetism and optics. Quantum theory unifies chemistry, properties of matter, and ultimately, with the development of quantum electrodynamics, electromagnetic phenomena. General relativity unifies special relativity, gravitation and the structure of space-time. Quantum electroweak theory (partially) unifies the electromagnetic and weak forces. The so-called standard model (partially) unifies all known phenomena apart from gravitation. String theory, or M theory, if successful, will unify all phenomena. The very phenomenon that Kuhn holds to mark discontinuity, namely revolution, actually also reveals continuity – continuity of the search for, and the successful discovery of, underlying theoretical unity. Revolutionary developments in theoretical physics all reveal one common theme: the increasingly successful capture of physicalism, more and more adequately, as a single, precise, unified, testable, physical "theory of everything". Almost the whole of theoretical physics since Galileo substantiates the claim that physicalism is by far the most fruitful idea that science has come up with at that level in the hierarchy of assumptions. The whole way in which theoretical physics has developed points at physicalism.

But in order to appreciate this point, it is essential to adopt a generalized hierarchical view, of which AOE is a special case. The historical record may reveal discontinuity at levels 2 and 3; we need to recognize level 4 to appreciate continuity through this lower-level discontinuity. Indeed, if we take pre-Galilean, Aristotelian science into account, we would need to ascend to level 5 to see continuity through the Galilean revolution. As a result of restricting himself to levels 1 and 2, and perhaps level 3, Kuhn was unable to discern theoretical continuity across the discontinuity of revolutions (although this is manifest, even for Kuhn, in revolutions other than those in theoretical physics).

One criticism that may be levelled against AOE is that it does not just accurately *reflect* scientific practice, but has the audacity to claim to *correct* scientific practice. It does this by providing a framework for the articulation and scrutiny of level 3 metaphysical blueprints, as an integral part of science itself, thus providing a rational, if fallible, means for the development of new non-

empirical methods, new symmetry principles, and new theories. AOE makes explicit what is at present only implicit, due to the current influence of SE on the scientific community. And more generally, AOE has implications for scientific practice throughout the natural sciences in depicting scientific method in a hierarchical, meta-methodological fashion. Does this not tell against AOE? No. Any new conception of science which substantially improves our understanding of science ought to enable us to improve scientific practice. It would be very odd if our ability to do science well were wholly divorced from our understanding of what we are doing. A test for a new theory of scientific method ought to be, then, that it improves scientific practice, and does not merely accurately depict current practice. AOE passes this test.

In case it should seem miraculous that science has made progress without AOE being generally understood and accepted, I should add that good science has always put something close to AOE into practice in an implicit, somewhat covert way, and it is this which has made progress possible.

3 The Model of the Aesthetic Induction

I turn now to James McAllister's account of the role of non-empirical, aesthetic factors in the selection of theories in science. I shall call McAllister's view "the model of the aesthetic induction" (MAI). Here, in summary, is his view.

According to MAI, the basic aim of science is to develop a body of theory that successfully predicts all observable phenomena. MAI holds that from this aim of "empirical adequacy", we can arrive at the following criteria for assessing theories: success in predicting existing empirical data, success in predicting new phenomena, consistency with other high-level theories, explanatory power, and internal consistency.

Many scientists have however declared that aesthetic considerations, in addition to the above, play a vital role in both the discovery and acceptance of theories in science. Dirac, Einstein, and many others have stressed the importance of aesthetic considerations, such as beauty, elegance, harmony, uniformity amidst variety, simplicity, symmetry. MAI holds that such criteria do indeed have an important role to play in deciding what theories

are accepted, to the extent, even, on occasions, of over-riding empirical considerations.

But, according to MAI, in so far as such aesthetic considerations exercise a rational influence over choice of theory in science, two crucial points need to be borne in mind. First, theories must be considered to be abstract entities, distinct from this or that linguistic formulation. Second, what matters is not the (subjective) aesthetic judgements themselves, but rather objective, non-aesthetic properties that theories, construed as abstract entities, do actually possess, in virtue of which scientists make their aesthetic judgements.

There are, according to MAI, five classes of properties of theories that are relevant: symmetry, invocation of a model, visualizability/abstractness, metaphysical allegiance, and simplicity (related to unity). MAI stresses that many different properties fall under each of these headings. There are different kinds of symmetry; different theories have different kinds of models; some scientists, in some contexts, hold visualizability to be a virtue, while others, in other contexts, prize almost its opposite, namely abstractness; scientists have upheld different metaphysical views at different stages in the development of science, in terms of which they have sought to interpret scientific theories; and there are many different ways of assessing the simplicity of theories, yielding quite different results.

How, then, does the scientific community decide which of these very many different kinds of properties of theories are the relevant or important one's to employ in order to assess the acceptability of theories on non-empirical, or aesthetic grounds? And what is the justification for so assessing theories, in terms of the preferred properties? How, in particular, can MAI do justice to the fact that aesthetic criteria in science change over time?

The answer is that, at any given stage, a scientific community prefers those new theories which have properties which earlier theories, which have proved to be empirically successful, also possess. If a certain kind of theory, with characteristic aesthetic properties, has met with empirical success in the past then, understandably enough, scientists are influenced to give preference to similar kinds of theories in the future, with similar properties.

This is "the aesthetic induction". At a stroke, the above three questions are answered.

In a little more detail, we can imagine that a scientific community can consider many different aesthetic properties of theories, P, Q, R,... The community will assign a different weighting, W_P , W_Q , W_R , to each of these properties, each weighting determining how influential the corresponding property is in theory choice. The weightings are in turn determined by what kinds of theory, with what properties, have (or have not) met with empirical success in the past. W_P , W_Q , W_R ... are, in other words, determined by the aesthetic induction.

According to MAI, then, two kinds of criteria are employed in science to choose theories. On the one hand there are criteria, listed above, arrived at by analysis of the basic aim of science of achieving empirical adequacy. And on the other hand, there are criteria arrived at by the aesthetic induction. The second presupposes the first.

Aesthetic criteria will tend to be conservative, based as they are on empirical performance of theories in the past. New theories, with the potential for great predictive success, may violate existing, conservative aesthetic criteria. When such a theory is developed, there is a rupture in accepted aesthetic criteria. Initially the new theory is judged to be "ugly"; but as its empirical potential becomes manifest, aesthetic criteria are changed to suit the new theory. This is what a scientific revolution amounts to, according to MAI, a conception somewhat different from Kuhn's. In terms of this conception, neither Copernicus's theory, nor Einstein's theory of special relativity, were revolutionary, because neither broke with aesthetic criteria of the past. But Kepler's laws of planetary motion, and quantum theory, were both revolutionary, in that these theories broke dramatically with aesthetic criteria generally accepted at the time.

Finally, though the aesthetic induction might one day favour some particular metaphysical world view, so far this has not happened (McAllister, 1996, 102-4).

4 Comparison of the Two Views

What is rather astonishing about AOE and MAI is that, though arrived at independently, and though giving what are, in some respects, very different pictures of the scientific enterprise, nevertheless the two views have much in common. Both seek to uphold what McAllister calls "the rationalist image of science". Both hold that (some) criteria of theory choice can be justified by an appeal to the aims of science. Both hold that non-empirical criteria of theory choice have an enormously important part to play in science. Both hold that these non-empirical criteria are, in practical applications, quite diverse in character. Both hold that they change over time, as science progresses. And there is considerable agreement as to what these non-empirical criteria are: simplicity, unity, symmetry, and compatibility with some metaphysical world view, are all important, for both views.

Both hold that these criteria apply, not to specific formulations of theories, but to what all possible formulations have in common. And both give accounts of scientific revolutions that differ substantially from Kuhn's account.

But there are also dramatic differences. MAI is, for McAllister, "a medium-level model of scientific practice, of a scope intermediate between the loftiest generalization and the historical study" (1996, 2). AOE is put forward as a "highest-level model", with implications and applications for all of natural science. (Strictly speaking, it is what I call "generalized AOE" [Maxwell, 1998, 191-2, 185, 191, 208 and 223-4], embodying the hierarchical structure of AOE, but lacking specific, lower level theses of AOE, that is a model at the highest level, applying to science throughout history; AOE is restricted to post-Galilean science.)

Again, MAI is a version of SE, whereas AOE emphatically rejects SE. That MAI is a version of SE is clear from the way the aim of science is characterized as "empirical adequacy". It is also apparent in the way science can, according to MAI, establish a metaphysical world view. This can only happen via the aesthetic induction, and has not as yet come about. According to AOE, by contrast, at levels 7 and 6 there are metaphysical, cosmological theses that are permanently accepted by science, and at levels 5 to 3, there are metaphysical theses which are a part of current scientific knowledge, but which are increasingly likely to require

revision with the advance of science, as one descends from level 5 to level 3: see diagram 2.

Whereas MAI gives to science just one aim (empirical adequacy), AOE sees science as having a hierarchy of aims, from empirical adequacy, perhaps, at the highest level, down to the aim to turn the best available level 3 blueprint into a precise, true theory of everything, at the lowest level. (And even more specific, and different, aims are assigned to different branches of natural science.)

That AOE postulates this hierarchical structure to the aims of science, whereas MAI does not, leads to different treatments of changing criteria for theory choice. According to MAI, criteria of theory choice are of two kinds: those that are justified by an appeal to the basic aim of science (empirical adequacy), and those that are justified by inductive projection – the aesthetic criteria arrived at by the aesthetic induction. These latter are weaker than the former, and presuppose, for logical reasons, the former (McAllister, 1996, 76). According to AOE, by contrast, all criteria of theory choice are arrived at by aim analysis: those that evolve do so because the level 3 aim of science evolves.

Even though MAI and AOE agree that non-empirical criteria of theory choice change with time, they disagree about what criteria change, and what this change involves. According to AOE, something close to physicalism has been implicit in the methods of theoretical physics since Galileo or Newton; the demand for theoretical unity, associated with physicalism, has been more or less unchanging. What has changed is the form that the demand for unity takes, as manifest in dramatically changing level 3 metaphysical blueprints. MAI does not claim that physicalism, and the requirement of unity associated with it, is a part of the unchanging criteria of theory choice (since Galileo, at least). Nor could MAI claim this, as long as it is a version of SE.

According to AOE, the level 4 thesis of physicalism, and the level 3 best metaphysical blueprint, are arrived at by a quasi-Popperian process of conjecture and criticism, the whole direction of progress in theoretical physics being taken into account since the birth of modern science (or since the Presocratics). The claim is that these theses make explicit what theoretical physics hopes to

achieve: they are intended to be the best conjectures as to what the basic aims of theoretical physics should be, at different levels of specificity. These theses are intended to lead to criteria, to methodological principles such as symmetry principles, that will be relevant for future theories, not yet developed. Indeed, according to AOE, the activity of further articulating the best blueprint, and solving problems of unity to which it gives rise, provides science with a rational, if fallible method of discovery. All this contrasts dramatically with criteria arrived at by the aesthetic induction, according to MAI, which are almost bound to be conservative, and more or less inapplicable to revolutionary developments. AOE criteria anticipate and provoke revolution, and judge the existing body of fundamental physical theory as unsatisfactory because of its failure to comply with the demand for unity (the standard model postulates too many particles and forces, and clashes with general relativity); by contrast, MAI criteria are conservative, and are almost bound to be at odds with revolutionary developments (McAllister, 1996, 81-5 and 128-33). AOE criteria are heuristically powerful; MAI criteria are the opposite. Furthermore, AOE criteria, associated with level 3 blueprints, evolve or improve as physics makes progress, and in a way which admits some elements of continuity: see, in particular, Maxwell (1998), 80-9. MAI sees change, but no overall progress, in non-empirical criteria, and holds, in a quasi-Kuhnian fashion, that revolutions create a rupture in aesthetic criteria, there being no account of the modification and generalization of blueprints, which AOE provides.

MAI and AOE agree that non-empirical criteria apply, not to any specific formulation of a theory, but to what all formulations have in common. But there are somewhat different accounts of what this is. According to MAI, a formulation-independent theory is an abstract entity that exists in its own right, with its own properties distinct from the phenomena the theory postulates (see, for example, McAllister, 1996, 98-100). This leaves obscure what sort of thing such an abstract entity is, and what its relationship is with a linguistic formulation of the theory, and with the phenomena it predicts. According to AOE, the matter is much more straightforward: a formulation-independent theory, T , is the

content of T, what T predicts, or asserts to be the case. AOE does not appeal to the abstract entities of MAI; it appeals only to possible phenomena, not as actually existing entities, but merely as possibilities. The claim that T exhibits a certain symmetry thus amounts to the claim that the phenomena predicted by T exhibit this symmetry. There is here no mystery about the relationship between a linguistic formulation of T, the abstract entity T, and the phenomena T predicts: the "abstract entity" is just what any linguistic formulation of T asserts to be the case, the *content* of T. This leads to an account of the importance of linguistic-dependent criteria of simplicity (Maxwell, 1998, 110-3), something which MAI does not provide.

A fundamental difference between MAI and AOE, encapsulated in the title of this chapter, is that, whereas MAI holds that aesthetic criteria are important in science, AOE denies this, all non-empirical criteria for theory choice being reducible to the demand that the totality of fundamental physical theory exemplify the level 4 thesis of physicalism or, more specifically, the best available blueprint at level 3. For AOE, what matters is unity or comprehensibility, not beauty.

But this difference is terminological rather than substantial. McAllister defends a projectivist, subjectivist account of the beauty of theories. Beauty is in the eye of the beholder, rather than in the theory itself.²³² Scientists judge certain theories to be beautiful because of non-aesthetic properties that they possess objectively; it is these non-aesthetic properties that are important methodologically and epistemologically, and play the crucial role in the aesthetic induction. One of these is metaphysical allegiance. The demand that the totality of fundamental physical theory should exemplify physicalism, and the best available blueprint, are special cases of metaphysical allegiance. Comprehensibility is just one of McAllister's aesthetic properties. Comprehensibility, one might say, is beautiful. It fits perfectly Hutchinson's characterization of beauty (McAllister, 1996, 17-23) as involving "uniformity amidst

²³². Elsewhere I have defended an objectivist, realist account of value: see Maxwell (1984) chapter 10, (1999b) and (2001a) chapter 2. This does not, however, affect the present argument.

variety": see the discussion of "unity through diversity" in Maxwell (1998), chapter 3.

A more serious disagreement would seem to be that whereas AOE recognizes only *one* methodologically significant non-empirical property, namely unity or comprehensibility, MAI stresses that there are endlessly many, falling under the five headings of symmetry, invocation of a model, visualizability/abstractness, metaphysical allegiance, and simplicity (related to unity).

This disagreement is not quite as big as it might at first appear to be. Here, very briefly, are the similar, but also different, ways in which AOE and MAI treat unity, symmetry, metaphysical allegiance and simplicity.

Unity. AOE and MAI both recognize that the demand for unity takes a number of different forms, but AOE alone holds that these are aspects of just one, single conception of unity. According to AOE, dynamic unity, postulated to exist by physicalism, can be broken in thought in a number of different ways, this creating a number of different kinds of (relative) disunity, and hence a number of different ways in which degrees of unity (or disunity) can be assessed. But these different kinds of disunity all relate to just one conception of unity, namely that which is postulated to exist by physicalism: see Maxwell (1998), 89-93, and 280 note 22. MAI too stresses that there are different kinds of unification (McAllister, 1996, 110) but, unlike AOE, does not relate these to one basic conception of unity.

Symmetry. Here, again, AOE and MAI both recognize that the demand for symmetry takes a number of different forms, but AOE alone holds that these, in so far as they are methodologically legitimate within theoretical physics, all relate to the one basic demand for unity. One of the achievements of AOE is to demonstrate clearly how different kinds of symmetry relate to unity, the demand that theories exhibit symmetries itself being an aspect of the demand for unity (Maxwell, 1998, 89-103, 123-40 and 257-65). MAI recognizes that theories exhibit different kinds of symmetry (McAllister, 1996, 41-4), but fails to recognize that different kinds of symmetry, in theoretical physics at least, are aspects of unity.

Metaphysical Allegiance. Once again, both AOE and MAI recognize that an important non-empirical requirement in theoretical physics, upheld by some physicists at least, is that fundamental physical theories should accord with some metaphysical view. Both recognize that metaphysical views associated with physics since Galileo have changed dramatically over time; both recognize that different physicists espouse different metaphysical views at the same time. But AOE and MAI differ, here too, in that AOE holds that diverse, evolving level 3 blueprints, in order to be acceptable, need to accord with physicalism, whereas MAI makes no such demand. For AOE, the requirement that a theory exemplifies a metaphysical view, in so far as it is methodologically legitimate, is but an aspect of the basic requirement that the body of fundamental physical theory exemplifies the unity of physicalism (as much as possible). MAI makes no such demand.

Simplicity. Here, yet again, both AOE and MAI recognize that the demand for simplicity takes a number of different forms; both see simplicity as being related to unity, but in somewhat different ways: compare Maxwell, 1998, 111-3 and 157-9, with McAllister, 1996, 109-11. But AOE alone relates the demand for simplicity to the more basic demand for just one kind of unity, dynamic unity postulated by physicalism.

The difference that this reveals in the two views can be summed up like this. AOE postulates just one basic non-empirical requirement, unity, and relates different requirements, of different kinds of unity, symmetry, metaphysical allegiance and simplicity, to this one demand for unity. MAI, by contrast, holds that there just are many different kinds of requirements of different kinds of unity, symmetry, metaphysical allegiance and simplicity. Unlike AOE, MAI sees no unity in diverse kinds of unity, symmetry, metaphysics and simplicity. (In this respect, AOE might be said to give a more unified, and hence more beautiful, account of scientific method than MAI.)

More substantial differences arise in connection with the two remaining kinds of aesthetic properties of theories which MAI holds to be methodologically significant, which I now consider in turn.

Invocation of a Model. AOE recognizes that an important consideration in assessing a new physical theory is that it has a form similar to existing empirically successful physical theories. Thus the acceptability of quantum electroweak theory, and chromodynamics is much helped by the fact that these theories are similar in form to the highly empirically successful theory of quantum electrodynamics. All three theories, despite their differences, are locally gauge invariant quantum field theories. According to AOE, this requirement of similarity of form or structure derives, once again, from the requirement of unity (Maxwell, 1998, 112). If T_1 and T_2 have some similar structure, then some part of T_1 can be modelled by some part of T_2 , and vice versa. According to AOE, having a model is only methodologically significant to this extent, and once again this requirement turns out to be derived from the demand for unity. (Of course, that physical reality is a model of a theory, in the sense that the theory is *true*, is highly significant for AOE; but this is not what MAI means by a "model".) MAI is, once again, much more open-ended in the kind of models that it is prepared to recognize as methodologically significant, and does not attempt to derive the requirement that a theory should have some kind of model from the demand for unity.

Visualizability/abstractness. According to AOE, neither visualizability nor abstractness are methodologically significant for theoretical physics. What does matter is that a theory can at least be interpreted realistically, as postulating that such and such a physical entity, (or entities), such as a field (or particles) exists, a stepping stone towards the ubiquitous, unified *something* of physicalism. (Actually, AOE demands more. It demands that fundamental physical theories are open to being interpreted in terms of conjectural essentialism: see Maxwell [1998], 141-55.) If one has acquired an intuitive understanding of a realistic theory, then one may well be able to "visualize" what the theory is about: to this extent, visualizability is methodologically significant, according to AOE, but once again derives from the demand for unity, via the demand for realism. MAI, by contrast, once again, is much more open-ended about visualizability, and makes no attempt to relate it to the demand for unity.

McAllister claims that opposition to orthodox quantum theory (OQT), by Schrödinger, Einstein and others, stemmed from the loss of visualizability and determinism associated with the new theory. But this overlooks the key, entirely legitimate objection to OQT, namely its loss of microrealism, due to the failure to solve the quantum wave/particle problem. Because it failed to specify a consistent quantum ontology, OQT had to be developed as a theory which can, at most, make predictions about the results of performing measurements on quantum systems – measurement being described classically. But this in turn meant that OQT is, quite essentially, made up of two quite different parts stitched together in a grossly *ad hoc* way, namely (1) the quantum part, and (2) some part of classical physics for a treatment of measurement. Despite its immense empirical success, OQT is still today deeply and genuinely problematic, to the point, almost, of being unacceptable, because of its grossly *ad hoc* character, due to its lack of microrealism (Maxwell, 1972b, 1976b, 1982, 1988, 1998 chapter 7). The mature Einstein was well aware that this is the basic objection to OQT, not lack of visualizability or loss of determinism (Maxwell, 1993a, 290-6). Elsewhere I have argued that the grossly *ad hoc* character of OQT, stemming from its lack of microrealism, provides us with a general argument against instrumentalism and for realism (Maxwell, 1993b). I have also suggested how the quantum wave/particle problem may be solved, and how a fully microrealistic version of quantum theory may be developed, free of any reference to measurement or classical physics in its basic postulates, able to recover all the successful predictive content of OQT, but also making experimental predictions different from OQT for as yet unperformed experiments (Maxwell, 1976, 1982, 1988, 1998 chapter 7, and especially 1994). (This was done in an attempt to put the rational, but fallible, method of discovery of AOE into scientific practice.) There are, of course, other attempts at developing fully microrealistic versions of quantum theory: see Bohm (1952), Ghirardi, Rimini and Weber (1986), and Penrose (1986).

Einstein's mature objection to OQT had to do with the lack of realism of the theory, but he did also, especially earlier, object to its lack of determinism. But here, too, there is a methodologically

significant issue at stake, related once again to the demand for unity. A realistic version of quantum theory must be unified, first with special relativity, and then, ultimately, with general relativity. This is a much graver problem, granted probabilistic quantum theory, than it is if quantum theory is deterministic. The demand for unity speaks against probabilistic quantum theory – but not decisively: nature may well be probabilistic, and the task may be to develop probabilistic versions of special and general relativity (Maxwell, 1985a).

As for abstractness, this is, for AOE, without methodological significance except that, as physical theory draws closer to capturing physicalism, it is almost bound to specify entities increasingly remote from those of ordinary experience. We begin with corpuscles, minute billiard balls, in the 17th century; these then transmute into point-particles that interact by means of forces; these, in turn, transmute into classical fields, into quantum fields, into curved space-time, into superstrings in ten dimensional space-time – entities increasingly remote from the familiar billiard ball.

We have seen, so far, that AOE recognizes, ultimately, just one non-empirical criterion, unity or compatibility with physicalism, whereas MAI recognizes many, and makes no attempt to show that these all devolve from just one basic criterion. But I come now to a non-empirical criterion which AOE holds to be absolutely central, but which MAI does not even recognize as an aesthetic criterion at all: explanatory power.

Explanatory power is an ambiguous concept. We may hold that T_1 has more explanatory power than T_2 if (1) T_1 has greater empirical content than T_2 , or if (2) T_1 has greater unity than T_2 even though the same empirical content. Let us call these type (1) and type (2) explanatory power respectively.

We need also to recognize that criteria legitimately employed in science to assess theories can be put into three categories: (a) empirical, (b) empirical-dependent, and (c) non-empirical. By (a) I mean simply the predictive success of the theory in question; by (b) I mean properties of theories that have to do with how amenable they are to being assessed empirically, such as testability and empirical content; and by (c) I mean properties of theories that

have nothing directly to do with empirical success but which are deemed to indicative of truth, or of potential empirical success.

Type (1) explanatory power is a typical type (b) property of theories. But, according to AOE, type (2) explanatory power is the key type (c) non-empirical property of theories from which, as we have seen, all others, such as symmetry, simplicity or metaphysical allegiance arise. In seeking to acquire knowledge about the world, we actively hunt for clues as to the kind of universe we are in, and hence the kind of theories we need to develop. *The* big clue that we have (apparently) discovered is that the universe is more or less comprehensible in some way or other, it being possible to discover explanations for phenomena; this is enshrined in theses of near and rough comprehensibility, at levels 6 and 7 of figure 1 on page 8 of Maxwell (1008). We then make the bold conjecture, at level 5, that the universe is perfectly comprehensible in some way or other – the universe being such that there is some one kind of explanation for all phenomena, couched in terms of God, a cosmic purpose (which everything is designed to fulfil), a cosmic programme, a unified physical entity, or something else. From Galileo on, science has, in effect, made the even bolder conjecture that the universe is *physically* comprehensible, at level 4, and comprehensible in terms of the best available blueprint, at level 3. Type (2) explanatory power, to repeat, is the *key* type (c) non-empirical criterion of theory choice, from which all other type (c) criteria arise. If any property of theories cried out to be *the* aesthetic property of beauty, which scientists quite properly take note of as being methodologically significant, it is type (2) explanatory power.

And yet, astonishingly, MAI does not even include explanatory power in its list of aesthetic properties of theories, despite its open-ended, all-inclusive approach to listing such properties (in such sharp contrast to AOE).

MAI holds that the requirement of type (2) explanatory power is a permanent criterion of science, one which can be arrived at by aim-analysis, taking the aim of science to be empirical adequacy (McAllister, 1996, 11). It is clear that type (2), and not merely type (1) explanatory power is intended here, for McAllister writes that a successful explanatory theory is deemed to have "identified a

pattern or mechanism underlying the data" (1996, 11). But such an analysis could only, at most, justify adopting the requirement of type (1) explanatory power; it does not justify adopting type (2) explanatory power as a requirement – not unless the truth, the universe that is, is permanently presumed to have a more or less unified dynamic structure (a presumption which contradicts SE). McAllister provides no argument in support of the contention that favouring theories with type (2) explanatory power can be justified by an appeal to the aim of empirical adequacy. He does refer to an approach to the problem of induction, espoused by Braithwaite and Mellor, according to which we are justified in proceeding as if regularities or patterns exist in nature because this gives us the best hope of acquiring knowledge whatever the universe may be like (McAllister, 1996, 100-1). It is this argument, perhaps, which McAllister assumes justifies taking type (2) explanatory power as a permanent criterion for theory choice, arrived at by aim-analysis, taking the aim of science to be empirical adequacy.

But there are three things wrong with this.

First, the Braithwaite-Mellor justification of induction does not work, as I shall show in the next section.

Second, many different kinds of explanation are possible; the universe may be comprehensible (phenomena being explainable) in many different ways, and to many different degrees, as the different theses from levels 3 to 7 of AOE attest. Here, above all, science needs to be flexible and responsive, constantly modifying the kind of explanations to be sought in the light of empirical success and failure, in the kind of way in which the hierarchical methodological structure of AOE is designed to facilitate. If ever there was a role for the aesthetic induction, it would surely be here, in connection with explanatory power. But in excluding type (2) explanatory power from the list of aesthetic properties, and in making it a fixed, unchanging requirement of theory choice, MAI fails to exploit this vital need for science constantly to modify and improve the kind of explanations that it seeks. It is just this, by contrast, that is the key idea behind AOE.

Third, if McAllister's argument were successful, so that giving preference theories that exhibit type (2) explanatory power could be justified by an appeal to the aim of empirical adequacy, then

this would be a disaster for MAI, for it would obviate entirely the need for science to consider aesthetic properties of theories, and to employ the aesthetic induction. As I have argued above, all aesthetic properties of theories that have any methodological significance can be derived from the demand for unity – that is, the demand for type (2) explanatory power. Once type (2) explanatory power is acknowledged to be methodologically significant, no other aesthetic properties of theories are required by science.

I conclude this section by just mentioning three further differences between AOE and MAI.

First, reasons given in defence of MAI for holding that aesthetic considerations are methodologically significant in science arise from the fact that scientists themselves have stressed their importance, and they do indeed seem influence what theories are chosen in science. Reasons given in defence of AOE for holding that type (c) non-empirical considerations are methodologically important are much stronger: science becomes impossible if such considerations are not deployed to rule out endlessly many empirically successful but grossly *ad hoc* theories.

Second, MAI, despite being a contribution to the rationalist conception of science, does not provide a basis for systematically correcting scientific practice. But AOE does. As I have already remarked, if a view genuinely increases our understanding of science, it would be surprising if it did not have implications for scientific practice. AOE passes this test, in emphasizing the need for explicit articulation of metaphysical theses at levels 3 and 4, and explicit tackling of the problems thereby generated.

Third, MAI and AOE conceive of the relationship between science and the philosophy of science differently. MAI takes the conventional view for granted: philosophy of science is a meta-discipline which seeks to spell out and justify methods implicit in successful scientific practice, but which is quite distinct from science itself. AOE upholds the unorthodox view that the philosophy of science is an integral part of science itself, influenced by and seeking to influence science, articulating and critically assessing actual and possible aims and methods for science, at various levels, the fundamental aim being to contribute

to scientific progress. A new level 3 aim for physics, i.e. a new blueprint, plus associated new methods, might constitute a major contribution to theoretical physics, as well as being a contribution to the philosophy of physics. Einstein's special theory of relativity is an example. It puts forward both a modified blueprint (Newtonian space-time becoming Minkowskian space-time), and modified methods (Galilean invariance becoming Lorentz invariance): it is thus a major contribution to physics itself which is also a contribution to the aims and methods of physics – i.e to the philosophy of physics.

5 Assessment

Which is to be preferred, AOE or MAI? The two views need not, of course, be regarded as rivals. AOE is a highest level model, whereas MAI is a medium level model; one could consider accommodating MAI within AOE. This would require, however, that MAI be modified quite extensively, as the previous section has shown.

Interpreting AOE and MAI as rival rationalist accounts of science, my chief criticism of MAI is that it is a version of SE, and thus suffers from the defects that all versions of SE suffer from. Given any empirically successful theory, T, there will always be endlessly many *ad hoc* rivals to T, even more empirically successful than T, which will never even be considered within science, let alone considered and rejected. In persistently rejecting such *ad hoc* rivals, even more empirically successful than T, science makes a persistent assumption about the nature of the universe. This contradicts SE; and contradicts MAI.²³³

²³³ McAllister might, of course, reject SE, and defend MAI in such a way that MAI acknowledges that science makes a substantial, permanent metaphysical assumption about the nature of the universe - namely that the universe is such that no *ad hoc* theory is true. But at once two major problems arise. What precisely is this assumption in view of the fact that there is no sharp distinction between the *ad hoc* and the non-*ad hoc*? What is the justification for making this assumption? In order to answer these questions satisfactorily, it is necessary to adopt AOE, which involves abandoning those parts of MAI which clash with AOE.

McAllister might seek to evade this conclusion by arguing, as he does in his book, that non-empirical, aesthetic criteria that rule out acceptance of empirically successful, *ad hoc* rival theories, are too diverse in character, too changeable over time, to amount to the implicit acceptance of any persistent assumption. But such an argument collapses the moment one takes radically *ad hoc* theories into account of the kind considered in section 2 above, and in Maxwell (1998), 47-54. Rejection (or rather complete neglect) of such radically *ad hoc* theories persists throughout revolutions and all changes in aesthetic fashions in science. The persistent rejection of such theories unquestionably commits science to making a substantial metaphysical assumption about the nature of the universe.

McAllister might, at this point, appeal to the pragmatic justification of induction of Braithwaite and Mellor, already referred to above (McAllister, 1996, 100-1). According to this argument, science proceeds, and is justified in proceeding, as if it assumes there are regularities to be discovered, but does not actually assume that regularities exist. But even if this argument is valid, it does not in any way invalidate my point above, that in persistently rejecting empirically successful *ad hoc* theories science implicitly makes a persistent metaphysical assumption about the world. It should be noted that kinds of *ad hoc* theories can be formulated that specify regularities, in that these theories are invariant with respect to position and time (no specific places or times being specified by the theories). These theories might be said to specify *ad hoc* regularities.

But in any case the Braithwaite-Mellor attempt at solving the problem of induction does not succeed. Restricting science to the search for regularities is both too narrow, and not narrow enough. Too narrow, because it is conceivable that we can live and acquire knowledge but not by searching for regularities in phenomena. God might get in touch with us, explain His purposes, keep us informed about what is going to happen. Getting in touch with God by means of prayer and meditation, and not by searching for regularities, might be the way to acquire knowledge; and various other science fiction possibilities can be imagined (see Maxwell, 1998, 185). Such possibilities are excluded by the search for

knowledge as characterized by Braithwaite and Mellor; this means these possibilities are just dogmatically assumed to be false. But the Braithwaite-Mellor approach is also not narrow enough because, as I have indicated above, if science is to be possible, *ad hoc* regularities must be persistently excluded from consideration. And, as we have seen, there is no sharp distinction between the *ad hoc* and the non-*ad hoc*. In previous chapters I have listed 8 kinds of disunity – in effect, 8 different ways in which regularities might be *ad hoc*, which range from the severely *ad hoc* (distinct regularities in different space-time regions) to the scarcely *ad hoc* at all (space-time and matter not being unified). What does the policy of "inductive projection" (McAllister, 1996, 101) amount to? Does it involve merely excluding permanently all theories that are type (1) *ad hoc* (distinct regularities in different space-time regions)? Again, this is both too narrow, and not narrow enough. Exactly the same objection arises wherever the line is drawn, between regularities that are too *ad hoc* to be considered by science, and those that are sufficiently non-*ad hoc* to be open to scientific consideration. We cannot, at this point, simply invoke the aesthetic induction, and declare that we discover, by induction, where the line is to be drawn between the acceptably and unacceptably *ad hoc*, because, as McAllister himself has so clearly shown, for logical reasons, the aesthetic induction can only proceed once methods have been arrived at by aim-analysis (McAllister, 1996, 76).

Another approach might be to favour permanently in science theories that are as non-*ad hoc* as possible, in all 8 senses, but not to draw a rigid line between the acceptably and unacceptably *ad hoc*. This would allow something like the aesthetic induction to proceed in science (although not in quite the open-ended way in which McAllister envisages). But even this attempt at solving the problem is both too narrow and not narrow enough. Endlessly many universes are possible in which we may live and acquire knowledge, and yet this inductive policy would not be appropriate for acquiring knowledge. It biases the search for knowledge in the direction of physicalism. But physicalism may be false; the universe may be comprehensible in some other way, or not comprehensible at all.

My claim is that the best possible way in which we can go about seeking knowledge is to do so employing the hierarchical aims-and-methods structure of (generalized) AOE. We must make some kind of guess as to what kind of universe we are in to proceed at all. At the top of the hierarchy we need to put those relatively contentless guesses which are such that their truth is required for acquisition of knowledge to be possible at all. These are justifiably permanent items of scientific knowledge. As we descend the hierarchy, we need to put increasingly contentful guesses, chosen because these seem to be the most fruitful from the standpoint of engendering methods that seem to offer the best help with acquiring empirical knowledge. As we proceed, we revise these guesses in the light of the relative empirical success and failure of rival research programmes, based on rival low-level metaphysical guesses. We try to keep such revisions as low as possible in the hierarchy when we seem to be achieving overall success, and only allow revisions to ascend higher up in the hierarchy when success is not being achieved, and higher-level revisions seem to be required.

This hierarchical conception of scientific method enables science to respond sensitively to what it seems to discover about the nature of the universe, lower-level aims and methods being adjusted in the light of apparent empirical success and failure, and within a framework of fixed, relatively unproblematic, higher-level aims and methods. All attempts at justifying induction pragmatically that are known to me, along lines advocated by Braithwaite and Mellor, fail because they fail to take note of the resources of (generalized) AOE. They all attempt to justify methods that are demonstrably not as efficient as those of AOE in enabling us to acquire knowledge of nature. They fail to encapsulate the responsiveness, the flexibility, the open-endedness and precision, of AOE.

And this is true of MAI as well. Indeed, as we saw above, in section 4, the aesthetic induction has conservatism built into it, and cannot help engender revolutionary new ideas for revolutionary new theories, whereas AOE is designed to do just that. It embodies a rational, if fallible, method of discovery for theoretical physics.

Chapter Twelve

A Mug's Game?

Solving the Problem of Induction with Metaphysical Presuppositions

This essay argues that a view of science, sketched here and expounded and defended in more detail elsewhere, solves the problem of induction. The view holds that we need to see science as accepting a hierarchy of metaphysical theses concerning the comprehensibility and knowability of the universe, these theses asserting less and less as we go up the hierarchy. It may seem that this view must suffer from vicious circularity, in so far as accepting physical theories is justified by an appeal to metaphysical theses in turn justified by the success of science. But this is rebutted. A thesis high up in the hierarchy asserts that the universe is such that the element of circularity, just indicated, is legitimate and justified, and not vicious. Acceptance of the thesis is in turn justified without appeal to the success of science. It may seem that the practical problem of induction can only be solved along these lines if there is a justification of the truth of the metaphysical theses in question. It is argued that this demand must be rejected as it stems from an irrational conception of science.

I

I think that I have solved a major philosophical problem: the problem of induction. This solution has been extremely fruitful, and it has enabled me to solve a good number of other philosophical problems. However, few philosophers would support the thesis that I have solved the problem of induction. Few philosophers have taken the trouble to study – or even to criticize – my views on this problem, or have taken notice of the fact that I have done some work on it.

This is how Karl Popper opens his book *Objective Knowledge*.²³⁴ There are at least two oddities about what Popper says here. First, Popper is wrong; he did not solve the problem of induction. Second, even by 1971, when this passage was first published,

²³⁴. Popper (1972, p. 1).

Popper's work on the problem of induction had received a great deal of attention.

Popper's words might, however, be uttered by me with far greater justice. For I really have solved the problem of induction. The solution has been extraordinarily fruitful, and has enabled me to solve a number of other philosophical problems.²³⁵ But few philosophers – if any – would agree that I have solved the problem. Few, indeed, have taken the trouble to study, or criticize, my work, or are even aware that I have done some work on the problem.²³⁶

I think I know why this is the case. First, it is no doubt the fate of most of us seeking to contribute to philosophy: our work sinks without trace, without comment. Second, the problem of induction has been around for a very long time; anyone claiming to solve the problem is almost bound to be wrong. Third, there is a kind of "negative judgement through persistent neglect" effect. The first version of my proposed solution was published nearly thirty years ago: if there was anything in it, surely someone would have noticed and taken up the idea, by now. Fourth, as Popper points out elsewhere,²³⁷ "analytic" philosophy has tended to be more interested in analysis of concepts than in proposed solutions to fundamental philosophical problems. Fifth, my solution amounts to a radical improvement of Popper's attempted solution. Popper was hostile to this, and Popperians today are hostile to it, precisely

²³⁵. Problems of unification and verisimilitude, the problem of rational scientific discovery, the problem of saying what it is that science has discovered about the ultimate nature of reality, problems concerning rationality, and the nature of social inquiry and, most important of all, the discovery that academic inquiry as it exists at present is profoundly defective when viewed from the standpoint of its capacity to help us learn how to become more civilized, there being an urgent need to bring about a revolution in the overall aims and methods of inquiry if we are to have what we so urgently need, a kind of inquiry rationally devoted to helping us acquire wisdom. See, for example, Maxwell: (1976; 1980; 1984a; 1984b; 1985b; 1987; 1991; 1992a; 1994b; 1997b; 1998 ch. 3, 4 and 6; 2000c; 2002d; 2003a; 2003b; 2003c; 2003d).

²³⁶. Work of mine related to my proposed solution to the problem of induction has received some critical attention: see for example: Kneller (1978), pp. 80-91; Longuet-Higgins (1984); Collingridge (1985); Midgley (1986); Easlea (1986); Smart (2000); Hodgson (2002); Muller (2004).

²³⁷. Popper (1963, ch. 2).

because I have the temerity to claim that I have radically improved Popper's ideas. Anti-Popperians are indifferent because they know Popper has failed to solve the problem, and they assume my approach inherits Popper's failure. Finally, and perhaps most damagingly, my proposed solution involves recognizing that science makes a persistent metaphysical assumption of "uniformity" or "unity". Philosophers at once know that any attempt to solve the problem of induction along these lines is hopeless. As Bas van Fraassen once put it "From Gravesande's axiom of the uniformity of nature in 1717 to Russell's postulates of human knowledge in 1948, this has been a mug's game" (van Fraassen, 1985, pp. 259-60). There is no need to study or criticize my proposed solution to the problem of induction: I am playing a well-known mug's game.

There is not much that I can do about the first five reasons for ignoring my work on the problem of induction: I can however at least set out to demolish the sixth reason. This is what I propose to do in what follows. I first give a brief sketch of my proposed solution to the problem of induction (spelled out in much greater detail elsewhere); I then demolish the thesis that it amounts to van Fraassen's "mug's game".²³⁸ My hope is that this may provoke one or two readers to take note of what I have done (see Maxwell, 1998; see also Maxwell: (1968, 1972a, 1974, 1976a, 1977a, 1979, 1980, 1984a, 1993a, 1997b, 1999a, 2000b, 2001a, 2002d, 2002e, 2003d).

II

My solution to the problem of induction is contained in a view about the aims and methods of science – a philosophy of science²³⁹

5. My earlier defense against the charge of circularity has been brief and unsatisfactory: see Maxwell (1998, pp. 166-8).

²³⁹. One of the great mistakes made by endlessly many attempts at solving the problem of induction is to assume unthinkingly that science is wholly in order just as it is, the task being to find some way of justifying existing valid methods of science. What the long-standing insolubility of the problem of induction is trying to tell us, in my view, is that the orthodox conception of science, taken for granted by scientists and non-scientists alike, is untenable, and needs to be changed. Before the problem of induction can be solved we need to *change* the currently accepted conception of science; indeed, we need to change, not just our conception of science, but science itself. Properly conceived, the problem of

– that I call "aim-oriented empiricism" (Maxwell, 1974, p. 140; 1998, pp. 6-13). In what follows I outline aim-oriented empiricism, indicate how it solves the “methodological” and “theoretical” problems of induction, demolish the thesis that aim-oriented empiricism represents a mug's game, and conclude by showing how the view solves what may be called the “practical” problem of induction.

The fundamental line of thought behind aim-oriented empiricism (AOE) can be indicated like this. Theoretical physics, and therefore all of natural science (since theoretical physics is fundamental for natural science), persistently selects fundamental physical theories that help to unify the whole of theoretical physics. Thus Newtonian theory (NT) unifies Galileo's laws of terrestrial motion and Kepler's laws of planetary motion (and much else besides). Maxwellian classical electrodynamics, (CEM), unifies electricity, magnetism and light (plus radio, infra red, ultra violet, X and gamma rays). Special relativity (SR) brings greater unity to CEM (in revealing that the way one divides up the electromagnetic field into the electric and magnetic fields depends on one's reference frame). SR is also a step towards unifying NT and CEM in that it transforms space and time so as to make CEM satisfy a basic principle fundamental to NT, namely the (restricted) principle of relativity. SR also brings about a unification of matter and energy, via the most famous equation of modern physics, $E = mc^2$, and partially unifies space and time into Minkowskian space-time. General relativity (GR) unifies space-time and gravitation, in that, according to GR, gravitation is no more than an effect of the curvature of space-time. Quantum theory (QM) and atomic theory unify a mass of phenomena having to do with the structure and properties of matter, and the way matter interacts with light.

induction involves formulating a new conception of the aims and methods of science, more rigorous than current conceptions, which is such that the problem of induction no longer arises. The task is not to justify the status quo, but to change the status quo so that the problem of justification no longer arises. Popper's failed attempt at solving the problem stands head and shoulders above the rest just because it fits this prescription: it consists of a new view about the aims and methods of science, a new philosophy of science, namely falsificationism.

Quantum electrodynamics unifies QM, CEM and SR. Quantum electroweak theory unifies (partially) electromagnetism and the weak force. Quantum chromodynamics brings unity to hadron physics (via quarks) and brings unity to the eight kinds of gluons of the strong force. The standard model unifies to a considerable extent all known phenomena associated with fundamental particles and the forces between them (apart from gravitation). The theory unifies to some extent its two component quantum field theories in that both are locally gauge invariant (the symmetry group being $U(1)\times SU(2)\times SU(3)$). String theory, or M-theory, holds out the hope of unifying all phenomena.²⁴⁰

It might be thought that, during the last 400 years or so, science has been pursued in a thoroughly open-minded, unbiased fashion, theories being selected impartially on the basis of empirical success alone, the emergence of increasing theoretical unity being a surprising and purely empirical discovery – unifying theories just being much more empirically successful than disunified rivals. Nothing could be further from the truth. In fact, in connection with every accepted unifying theory – NT, CEM, and the rest – there have always been endlessly many, easily formulated, disunified rival theories very much more empirically successful than the theories that have been accepted.²⁴¹

Thus, given NT, for example, one rival theory might assert: everything occurs as NT asserts up till midnight tonight when, abruptly, an inverse cube law of gravitation comes into operation. A second rival theory might assert: everything occurs as NT asserts, except for the case of any two solid gold spheres, each having a mass of a thousand tons, moving in otherwise empty space up to a mile apart, in which case the spheres attract each other by means of an inverse cube law of gravitation. A third rival asserts that everything occurs as NT asserts until three kilograms of gold dust and three kilograms of diamond dust are heated in a

²⁴⁰. For the AOE account of what theoretical unification means see Maxwell (1998, ch. 3 and 4).

²⁴¹. In what follows "accept T" implies, not just "accept T as a working hypothesis for further research", but also "accept T for the purposes of technological and other practical applications, including contexts where human life may depend on the predictions of T being true".

platinum flask to a temperature of 450°C, in which case gravitation will instantly become a repulsive force everywhere. And so on. There is no limit to the number of rivals to NT that can be concocted in this way, each of which has all the predictive success of NT as far as observed phenomena are concerned but which makes different predictions for some as yet unobserved phenomena.²⁴² Such theories can even be concocted which are *more* empirically successful than NT, by arbitrarily modifying NT, in just this entirely *ad hoc* fashion, so that the theories yield correct predictions where NT does not, as in the case of the orbit of Mercury for example (which very slightly conflicts with NT).²⁴³

This last point may be made more generally, as follows. Most accepted physical theories, for most of the time that they exist, are confronted by various empirical difficulties. Let T be any one of the above unifying accepted theories – NT, CEM, or whatever. Typically, T is confronted by the following empirical conditions. There is a domain A of phenomena for which the predictions of T are wholly successful; there is a domain B for which T fails to predict the phenomena because the equations of the theory cannot be solved; there is a domain C where T is ostensibly refuted (because the predictions of T clash with the phenomena of C, but this may be due, not to T yielding false predictions, but to experimental error, or relevant physical conditions not being taken into account); and finally there is a domain D of phenomena which T fails to predict because they lie outside the scope of T. (Here the phenomena, in A to D, are to be understood as consisting of low-level empirical or experimental laws.) It is now easy to concoct rivals to T that are much more empirically successful than T, as

²⁴². All the possible phenomena, predicted by any dynamical physical theory, T, may be represented by an imaginary "space", S, each point in S corresponding to a particular phenomenon, a particular kind of physical system evolving in time in the way predicted by T. In order to specify severely disunified rivals to T that fit all available evidence just as well as T does, all we need do is specify a region in S that consists of phenomena that have not been observed, and then replace the phenomenon predicted by T with anything we care to think of. Given any T, there will always be infinitely many such disunified rivals to T.

²⁴³. For a more detailed discussion of empirically successful *ad hoc* rivals to accepted theories, see Maxwell: (1974; 1993a; 1998, pp. 51-4).

follows. One such rival asserts: As far as phenomena in A are concerned, everything occurs as T asserts; as far as phenomena in B are concerned, the phenomena occur in accordance with established empirical laws; and the same for C, and for D. This rival to T, T* let us call it, reproduces all the empirical success of T (in A), successfully predicts phenomena that T is not able to predict (in B), successfully predicts phenomena that refute T (in C), and successfully predicts new phenomena, that lie beyond the predictive scope of T (in D). It might be demanded that T* should predict new phenomena; but this demand can be met too, since "phenomena" here, are laws with content in excess of actual experiments that have been performed. T* satisfies every imaginable requirement for being an empirically more successful theory than T.²⁴⁴

And this has been the situation for all the accepted fundamental physical theories indicated above, for most of the time that they have been in existence: endlessly many rival, disunified theories have been available, far more successful empirically than the accepted, unifying theories, and these empirically more successful, grossly disunified or, as I have called them "aberrant" theories (see Maxwell, 1974, p. 128) are all ignored.

As most physicists and philosophers of physics would accept, *two* criteria are employed in physics in deciding what theories to accept and reject: (1) empirical criteria, and (2) criteria that have to do with the simplicity and unifying capacity of the theories in question. (2) is absolutely indispensable, to such an extent that there are endlessly many theories empirically more successful than accepted theories, that lack unity, and are not even considered as a result.

Now comes the crucial point. In persistently accepting unifying theories, and excluding infinitely many empirically more successful, disunified or aberrant rival theories, science in effect makes a big assumption about the nature of the universe, to the

²⁴⁴. The two prescriptions for formulating empirically more successful rivals to accepted unifying theories, indicated in this and the previous paragraph, can of course be combined to create further more empirically successful disunified rival theories.

effect that it is such that no disunified theory is true, however empirically successful it may appear to be for a time. Furthermore, without some such big assumption as this, the empirical method of science collapses. Science would be drowned in an infinite ocean of empirically successful disunified theories.

If scientists only accepted theories that postulate atoms, and persistently rejected theories that postulate different basic physical entities, such as fields – even though many field theories can easily be, and have been, formulated which are even more empirically successful than the atomic theories – the implications would surely be quite clear. Scientists would in effect be assuming that the world is made up of atoms, all other possibilities being ruled out. The atomic assumption would be built into the way the scientific community accepts and rejects theories – built into the implicit *methods* of the community, methods which include: reject all theories that postulate entities other than atoms, whatever their empirical success might be. The scientific community would accept the assumption: the universe is such that no non-atomic theory is true.

Just the same holds for a scientific community which rejects all disunified or aberrant rivals to accepted theories, even though these rivals would be even more empirically successful if they were considered. Such a community in effect makes the assumption: the universe is such that no disunified theory is true.

Thus the idea that science has the aim of improving knowledge of factual truth, *nothing being presupposed about the nature of the universe independently of evidence* is untenable. Science makes one big, persistent assumption about the universe, namely that it is such that no disunified or aberrant theory is true. It assumes that the universe is such that there are no pockets of peculiarity, at specific times and places, or when specific conditions arise (gold spheres, gold and diamond dust, etc.), that lead to an abrupt change in laws that prevail elsewhere. Science assumes, in other words, that there is a kind of uniformity of physical laws throughout all phenomena, actual and possible. Furthermore, science *must* make this assumption (or some analogous assumption) if the empirical method of science is not to break down completely. The empirical method of science of assessing theories in the light of evidence can

only work if those infinitely many empirically successful but disunified theories are permanently excluded from science independently of, or rather in opposition to, empirical considerations; to do this is just to make the big, permanent assumption about the nature of the universe.²⁴⁵

Let us call this assumption of unity U; and let us call the view, just outlined, that in persistently only accepting unifying theories science presupposes U, "presuppositionism".

Most current views about science deny that science makes a substantial, persistent assumption about the universe. This is true, for example, of logical positivism, inductivism, logical empiricism, hypothetico-deductivism, conventionalism, constructive empiricism, pragmatism, realism, induction-to-the-best-explanationism, and the views of Popper Kuhn and Lakatos.²⁴⁶ All these views, diverse as they are in other respects, accept a thesis which may be called standard empiricism (SE): In science, theories are accepted on the basis of empirical success and failure, and on the basis of simplicity, unity or explanatoriness, but *no substantial thesis about the world is accepted permanently by science, as a part of scientific knowledge, independently of empirical considerations*. It deserves to be noted that even Feyerabend, and even social constructivist and relativist sociologists and historians of science uphold SE as the best available ideal of scientific rationality. *If science can be exhibited as rational, they hold (in effect), then this must be done in a way that is compatible with SE. The failure of science to live up to the rational ideal of SE is taken by them to demonstrate that science is not rational. That it is so taken demonstrates convincingly that they hold SE to be the only possible rational ideal for science (an ideal which cannot, it so happens, in their view, be met).*

²⁴⁵. For a much more detailed exposition of this refutation of standard empiricism see (Maxwell 1998, ch. 2).

²⁴⁶. For discussion of the claim that Kuhn and Lakatos defend versions of SE see Maxwell (1998), p. 40. Bayesianism might seem to reject SE, in acknowledging both prior and posteriori probabilities. But Bayesianism tries to conform to the spirit of SE as much as possible, by regarding prior probabilities as personal, subjective and non-rational, their role in theory choice being reduced as rapidly as possible by empirical testing: see Maxwell (1998), p. 44.

Presuppositionism is of course incompatible with SE, and thus incompatible with all the above doctrines. One crucial point needs to be noted about the argument so far: presuppositionism is more rigorous than all the above versions of SE *entirely independent of any justification of U, or justification for accepting U as a part of scientific knowledge* (that is in addition to the one given above). In saying this I am appealing to the following wholly uncontroversial requirement for rigour.

(R) In order to be rigorous, it is necessary that assumptions that are substantial, influential and problematic be made explicit – so that they can be criticized, so that alternatives may be developed and assessed.²⁴⁷

All versions of SE fail to satisfy (R) in just the way in which presuppositionism does satisfy (R). Presuppositionism makes the assumption U explicit (and so criticizable and, we may hope, improvable), while all versions of SE deny that science does make any such assumption as U. Thus presuppositionism is more rigorous than all versions of SE *even in the absence of any kind of justification of U*. In short, quite independent of any claim to solve the problem of induction, presuppositionism is more rigorous, and thus more acceptable, than any of the above versions of SE. This has a major implication for all attempts at solving the problem of induction: no such attempt can succeed if any version of SE is presupposed, since these all lack rigour. Attempts at solving the problem of induction must at least begin with presuppositionism, unless a better view of science emerges.

Far from presupposing the uniformity or unity of nature being a mug's game, it is all the other way round: attempting to construe science in such a way that science does *not* presuppose the uniformity or unity of nature is the mug's game, since all such views of science fail to satisfy elementary requirements for rigour, namely (R), and thus cannot provide a basis for solving the problem of induction that can hope to succeed. Presuppositionism is the only non-Mug's game in town unless, as I have said, something better turns up.

²⁴⁷. See Maxwell: (1984a), p. 224; (1998), p. 21.

Presuppositionism is, however, as it stands, untenable. This is because it is not at all clear what the assumption U is, or ought to be. It is vital to appreciate that there are endlessly many different assumptions of unity which science may be construed to make, almost all of which are false (since they contradict each other). Even more urgent than any problem of justification, there is the following problem: How can the assumption of unity being made by science at present, implicit in current scientific views as to what counts as theoretical unity, and almost bound to be false, be *improved*?

What is at issue is not the traditional philosophical problem of *justification* (which presupposes that U is true), but rather the scientific (and quasi-Popperian) one of *improving* what is almost bound to be false.²⁴⁸

²⁴⁸. What does it mean to say that U_2 is an "improvement" over U_1 , U_2 being "closer" to the true characterization of the unity of nature, U_t , supposing there is such a thing, than U_1 is? One answer is that U_2 is an improvement over U_1 if U_2 is such as to lead to a more empirically successful research programme than U_1 . But another answer can be given. Consider the steps that would need to be taken to move from U_t to U_2 and to U_1 . Suppose there is some proposition, U_{t2} , which is such that (1) U_{t2} is a true specification of the unity of nature that is less precise than U_t , so that $U_t \rightarrow U_{t2}$; (2) there is a way of making U_{t2} more precise so that the resulting proposition is U_2 , so that $U_2 \rightarrow U_{t2}$; (3) any proposition U^*_{t2} which satisfies (1) and (2) and is different from U_{t2} is less precise than U_{t2} (so that $U_{t2} \rightarrow U^*_{t2}$, but not vice versa). Suppose, too, that there is a proposition U_{t1} which satisfies (1), (2) and (3), but with U_2 replaced with U_1 . Then we can say that:

(A) U_2 is closer to U_t than U_1 is (thus being an improvement over U_1) iff $U_{t2} \rightarrow U_{t1}$ but not vice versa.

The intuitive idea, here, is that in order to get from U_t to U_2 , we need to make less drastic, less fundamental, modifications to U_t than we do to get from U_t to U_1 . U_2 is, as it were, more nearly a minor modification of U_t than U_1 is.

U_t , U_1 , U_2 , U_{t2} , etc. all need to be understood to be more or less precise versions of a true version of physicalism.

None of this has any epistemological import. Like the problem of verisimilitude, it is purely a problem of what it means to say " U_2 is closer to the truth than U_1 ".

It would also be possible to declare U_2 to be closer to the truth than U_1 if there are physical theories, T_2 and T_1 , precise versions of U_2 and U_1 respectively (so that $T_2 \rightarrow U_2$ and $T_1 \rightarrow U_1$, but not vice versa), such that T_2 is closer to the true theory of everything, T , than T_1 is (in the sense explicated in Maxwell, 1998, ch.

In surveying the different possible ways in which the universe may be unified, one important point to appreciate is that there is no single, sharp distinction between unity and disunity or aberrance. By "unity" we could mean merely that physical laws are the same throughout space and time. Or we could mean, in addition, that physical laws remain the same as other variables change, such as velocity, temperature, or mass (so that, for example, Newton's inverse square law of gravitation does not abruptly become an inverse cube law as masses of 1,000 tons are reached). Or, more restrictively still, we could mean (in addition) that there is only *one* force in nature, and not three or four distinct forces (such as gravitation, the electromagnetic force, and the weak and strong forces of nuclear physics). More restrictively still, we could mean that there is just one kind of particle in existence, or one kind of physical entity, a self-interacting field spread throughout space and time. Finally, and even more restrictively, we could mean that space, time, matter and force are all unified into one, unified, self-interacting entity.

Even more restrictive assumptions can be made, which specify the kind of entity or entities out of which everything is composed. And at the other end of the spectrum, much looser, less restrictive assumptions could be made which, if true, would still make science possible. Thus science could assume: the universe is such that local observable phenomena occur, most of the time, to a high degree of approximation, in accordance with some yet-to-be-discovered physical theory that is not too seriously disunified.

It is always possible, of course, that the universe only appears to be physically unified (to some extent). Perhaps, as theoretical physics advances, everything will become increasingly complex (as even some physicists believe²⁴⁹). Perhaps a malicious God is in charge, who has been controlling the universe up to now in such a way that it is as if physics prevails everywhere, but who, shortly, will startle us all by causing a series of dramatic, large-scale

5), and there is no other such pair of theories, T^*_2 and T^*_1 , such that T^*_1 is closer to T than T^*_2 is. This way of explicating " U_2 is closer to the truth than U_1 " is, however, likely to be applicable in a much more restricted way.

²⁴⁹. For example: Gordon Fleming (personal communication).

miracles to occur which violate all known laws. Perhaps as we probe deeper into physical reality we will discover that the universe exemplifies, not physical laws, but something that is closer to a computer programme (as some people have suggested). The universe may be comprehensible, but not *physically* comprehensible. That is, it may be that *something* exists – God, a society of gods, an overall cosmic purpose, a cosmic "computer" programme – which controls or determines the way events occur, and in terms of which, in principle, everything can be explained and understood: but this *something* may not be a unified physical entity, a unified pattern of physical law, and thus the universe, though comprehensible, is not *physically* comprehensible. Finally, the universe may not be comprehensible at all, and yet it might still be possible for us to live, and to acquire some knowledge of our local circumstances.

How do we choose between these endless possibilities? Science must make some kind of choice. It is all-important that science makes a correct choice, or at least a good choice, since this choice will determine what (non-empirical) methods are employed by science to assess theories. If science chooses a cosmological thesis that is radically false, then science will only consider false theories, and will exclude from consideration all theories that might take one towards the truth. Science will come to a dead end. The more restrictive the chosen cosmological assumption is, so the more potentially helpful it will be in selecting theories, but also the more likely the assumption is to be radically false, thus imposing a block on scientific progress. On the other hand, the looser, the more unrestrictive the assumption is, so the more likely it is to be true, but the less helpful it will be in excluding empirically successful "disunified" theories. (Other things being equal, the less one says, the more likely it is that what one says is true. "The universe is not a chicken" is almost certainly true about ultimate reality, just because it says so little, there being an awful lot of ways in which the universe can not be a chicken.)

It is all-important that science makes the right assumption about the ultimate nature of reality; and yet it is just here, concerning the ultimate nature of reality, that we are most ignorant, and are almost bound to get things wrong. How on earth are we to proceed?

The solution to this dilemma – the fundamental epistemological and methodological dilemma of science – is to make, not *one* cosmological assumption, but a *hierarchy* of assumptions, the assumptions becoming less and less restrictive, asserting less and less, as one goes up the hierarchy: see diagram 2 on page X, and associated text spelling out aspects of this hierarchical view.

The diagram makes things look complicated, but the basic idea is extremely simple. By displaying assumptions and associated methods – aims and methods – in this hierarchical fashion, we create a framework of high level, relatively unspecific, unproblematic, fixed assumptions and methods within which low level, much more specific, problematic assumptions and methods may be revised as science proceeds, in the light of the relative empirical success and failure of rival scientific research programmes to which rival assumptions lead. If currently adopted cosmological assumptions, and associated methods, fail to support the growth of empirical knowledge, or fail to do so as apparently successfully as rival assumptions and methods, then assumptions and associated methods are changed, at whatever level appears to be required.²⁵⁰ Every effort is made, however, to confine such

²⁵⁰. How can level 3 assumptions, or assumptions higher up in the hierarchy, both influence, and be influenced by, level 2 theories? What makes this possible is the feature of the hierarchy about to be indicated in the text, namely that, as one goes up the hierarchy, assumptions are more and more firmly upheld. Level 2 theories that accord with the best available level 3 assumption tend to be favoured over rivals that do not so accord. Nevertheless, a level 2 theory that clashes with the current level 3 assumption, but (a) accords with the level 4 assumption, and (b) is more empirically successful than theories that are in accord with the best level 3 assumption, will be accepted, and will lead to the rejection, or modification, of the level 3 assumption with which it clashes. Consider, however, a theory that clashes, not just with level 3, but level 4 as well, even though compatible with level 5, in such a way that no version of the idea that the universe is physically comprehensible, at level 4, can be rendered compatible with the theory. Such a theory would have to meet with far greater, sustained empirical success before it led to the overthrow of the current level 4 assumption. It would have to lead to empirical research programmes across a broad front of natural science even more successful than current science, based on the current level 4 assumption, before it would become acceptable. And this would be the case even more, given a theory that clashes with the current level 5 assumption. In short, an assumption at a given level may, for much of the time,

revisions to cosmological theses as low down in the hierarchy of theses as possible. Only persistent, long-term, dramatic failure (at levels 1 and 2) would lead us to revise ideas above level 3, let alone above level 4; only an earthquake in our understanding of the universe would lead us to revise ideas above level 5. In this way we give ourselves the best hope of making progress, of acquiring authentic knowledge, while at the same time minimizing the chances of being taken up the garden path, or being stuck in a cul de sac. The hope is that as we increase our knowledge about the world we improve the cosmological assumptions implicit in our methods, and thus in turn improve our methods. As a result of improving our knowledge we improve our knowledge about how to improve knowledge. Science adapts its own nature to what it learns about the nature of the universe, thus increasing its capacity to make progress in knowledge about the world – the methodological key to the astonishing, accelerating, explosive growth of scientific knowledge.

It is this conception of science, postulating more or less specific, problematic, evolving aims and methods for science within a framework of more general, relatively unproblematic, more or less fixed aims and methods, that I call *aim-oriented empiricism* (AOE).²⁵¹ [For further details see Maxwell (1998), chapters 1 and 3-6; 2004; and 2007, especially ch. 14.] The basic idea, let me re-

determine choices lower down in the hierarchy; but every now and again, it may itself be revised, because the revised version accords better with the assumption above, or is more empirically fruitful or, more likely, both of these simultaneously.

²⁵¹. Corresponding to each cosmological thesis, at level 3 to 7, there is a more or less problematic *aim* for theoretical physics: to specify that cosmological thesis as a true, precise, testable, experimentally confirmed "theory of everything". The aim corresponding to level 7 is relatively unproblematic: circumstances will never arise such that it would serve the interests of acquiring knowledge to revise this aim. As one descends the hierarchy of cosmological assumptions, the corresponding aims become increasingly problematic, increasingly likely to be unrealizable, just because the corresponding assumption becomes increasingly likely to be false. Whereas upper level aims and methods will not need revision, lower level aims and methods, especially those corresponding to level 3, will need to be revised as science advances. Thus lower level aims and methods evolve within the fixed framework of upper aims and methods.

emphasize, is that the fundamental aim of science of discovering how, and to what extent, the universe is comprehensible is deeply problematic; it is essential that we try to improve the aim, and associated methods, as we proceed, in the light of apparent success and failure. In order to do this in the best possible way we need to represent our aim at a number of levels, from the specific and problematic to the highly unspecific and unproblematic, thus creating a framework of fixed aims and meta-methods within which the (more or less specific, problematic) aims and methods of science may be progressively improved in the light of apparent empirical success and failure.²⁵²

This hierarchical view of AOE is put forward to solve the fundamental problem confronting presuppositionism, indicated above. It is put forward to solve the problem of *improving* the basic metaphysical assumption of science, implicit in persistent scientific preference for unifying theories even against the evidence, granted that some such assumption must be made, and it is almost bound to be false. The claim is that the hierarchical framework of AOE provides the best possible means for discovering metaphysical assumptions which best aid the task of improving knowledge of truth; AOE provides the best possible means for *improving* choice of metaphysical assumption. There is no attempt to justify the *truth* of metaphysical assumptions. At most, there is a justification for choosing metaphysical thesis A over B granted that the aim is to make that choice which gives the best promise of aiding the search for knowledge of truth. Justification is involved only in the quasi-Popperian sense that the

23. Some features of AOE may seem reminiscent of Laudan's "normative naturalism" (see Laudan, 1984, 1987). There are however marked differences: "normative naturalism" is not committed to physicalism, and does not postulate the hierarchy of aims and methods of AOE, which makes the rational assessment of low-level aims and methods possible. I might add that Laudan's "normative naturalism" is derived from AOE, which was first expounded in a Colloquium I gave at Pittsburgh University in 1972, chaired by Laudan, the text of which became Maxwell (1974). Rescher has defended the view that science makes metaphysical presuppositions (see Rescher, 1973, 1977, 1987); his views also differ substantially from AOE. For an excellent survey of recent methodological views, including those of Laudan and Rescher (but excluding AOE) see Nola and Sankey (2000).

best possible justification of metaphysical assumptions that we can have is to expose these assumptions to the most searching criticism possible, to criticism best designed to promote progress in knowledge.

Something like AOE has always been implicit in scientific practice (otherwise science would have come to an end). AOE becomes all but scientifically explicit with the work of Einstein in discovering special and general relativity. Aspects of this work that are characteristic of AOE are the fundamental role played by the search for theoretical unity, and the vital role played by symmetry principles (such as Lorentz invariance and the principle of equivalence). These latter are fallible and revisable, and have the dual role of being both physical and methodological principles, all of which is integral to AOE. (They are represented in the diagram by the slanting dotted lines.) For further details see Maxwell (1993a), pp. 275-305.

AOE, as indicated above, is intended to depict the metaphysical components of scientific knowledge given science as it exists today. AOE takes the specific form that it does in part because of what we have learned from Galileo onwards (or from the Presocratics onwards). History, in other words, is built into AOE. In the future, when we have learned more, AOE will be somewhat different. But however dramatic future revolutions in knowledge may be, we still ought to represent our knowledge in the same hierarchical form, with the same thesis, at level 7, at the top. Let us call this view "generalized AOE". When it comes to considering whether AOE succeeds in solving the problem of induction in a non-circular way, we need to consider various possible versions of generalized AOE which differ from AOE. The crucial question is: Can sufficiently good grounds be given for preferring AOE to all other rival versions of generalized AOE that one can think of? That is the proper way to formulate the problem of induction. (One striking feature of the problem of induction, as usually formulated, is its scientific sterility: work on the problem of induction has made no contribution to science, with the possible exception of Popper's work. But when the problem is formulated in the way just indicated, it is clear that it is potentially a highly fruitful problem for science: a version of generalized AOE that is

genuinely an improvement over AOE is likely to be a major contribution to science itself.)

III

At this point, the basic objection to the whole approach being advocated here may be reiterated. Either AOE solves the problem of induction, or it does not. If it does not, no more needs to be said. If it does, then an element of justification must enter in. This in turn means that AOE must commit van Fraassen's mug's game. Choice of theory, at level 2, is justified in part by being compatible with choice of metaphysical thesis at level 3 or 4; this latter choice is justified in turn in terms of the success of science. We have here the vicious circularity of the mug's game. And it is inescapable. Interpreting AOE as a framework for detecting error, for criticism, does not help; even given this interpretation, there must be some justification for regarding metaphysical thesis U_2 as a better choice, an improvement over, more likely to be true than, thesis U_1 : here, unavoidably, justification is present, which introduces the vicious circularity of the mug's game.

The first thing that needs to be said in response to this is that, as I have already emphasized, there is no question of justifying the *truth* (to some degree of certainty or probability) of any of the theses at levels 3 to 7. These theses remain, throughout, pure conjectures. I concur with Popper's thesis that all our knowledge is ultimately conjectural. (Whether such a view can claim to be the solution to the problem of induction is an issue I will take up below.)

At most, then, there is a justification for *accepting* such and such a thesis as a part of (conjectural) scientific knowledge, or *preferring* thesis A to thesis B.

Second, the top thesis is accepted on grounds which have nothing to do with the success of science at all. It is accepted because its truth is a necessary precondition for the acquisition of knowledge to be possible at all.

The thesis at level 7 asserts that the universe is such that it is possible for us to acquire some knowledge of our local circumstances (sufficient for it to be possible for us to continue to live). We are justified in accepting this thesis entirely in the absence of any justification for its truth (or probable truth), just

because we have nothing to lose; accepting this thesis as a part of our knowledge can only help, and cannot obstruct, the task of acquiring knowledge whatever the universe is like (see Maxwell, 1998, pp. 186-7).

This elementary argument for permanently accepting this level 7 thesis can of course be challenged. What is beyond question, however, is that no circularity is involved here at all. The argument in support of accepting the level 7 thesis makes no appeal to the success of science whatsoever. Science is not even mentioned.

I might add that a part of the point of exhibiting the metaphysical assumptions of science in the form of a hierarchy, from level 3 to 7, is to overcome a fatal objection to one traditional approach to solving the problem of induction, versions of which have been argued for by, for example, Reichenbach (1938, sections 38-41), Braithwaite (1953, pp. 255-92) and Mellor (1988). This argues that we are rationally entitled to assume that there are sufficient regularities in nature for the inductive methods of science to meet with success because, if such regularities do not exist, no method will procure knowledge. But this argument tries to establish too much; it is not valid. Counterexamples can be imagined. The world might be such that "the inductive methods of science" meet with no success at all, and yet we can still acquire sufficient knowledge to live (see Maxwell, 1998, p. 185). The thesis of AOE, at level 7, might be called a "principle of uniformity", but it is very much weaker than the assertion that there are regularities such that "the inductive methods of science" meet with success. The fatal objection to the Reichenbach-Braithwaite-Mellor (RBM) approach is that (1) *either* it seeks to justify acceptance of a "principle of regularity" which, if accepted, suffices to justify science, but the argument is invalid; (2) *or* it is valid, but the "principle of regularity or uniformity" whose acceptance is justified is much too weak to justify science. AOE adopts (2), and recognizes that the acceptance of other, more restrictive "principles of uniformity" needs to be justified on other grounds; RBM, not acknowledging the hierarchy of principles, are doomed to opt for (1). There is another, related objection to RBM: "the inductive methods of science", at least as conceived of by RBM, are *not* the

best available. They do not have the flexibility of the methods of AOE, which allow for the possibility of methods (associated with these low down in the hierarchy) being *improved* in the light of improving knowledge, feedback being facilitated by the hierarchical structure of AOE between improving knowledge and improving knowledge-about-how-to-improve-knowledge (i.e. improving aims and methods). The traditional "inductive methods of science", as a result of their inflexibility, are both too restrictive, and not restrictive enough. Like most other traditional attempts at solving the problem of induction, RBM try to justify the unrigorous, and thus the unjustifiable. The status quo needs to be changed, improved, not justified.

What about the thesis of "meta-knowability" at level 6?

Here are two arguments for accepting meta-knowability which make no appeal whatsoever to the success of AOE science.

(i) Granted that there is *some* kind of general feature of the universe which makes it possible to acquire knowledge of our local environment (as guaranteed by the thesis at level 7), it is reasonable to suppose that we do not know all that there is to be known about what the *nature* of this general feature is. It is reasonable to suppose, in other words, that we can improve our knowledge about the nature of this general feature, thus improving methods for the improvement of knowledge. Not to suppose this is to assume, arrogantly, that we already know all that there is to be known about how to acquire new knowledge. Granted that learning is possible (as guaranteed by the level 7 thesis), it is reasonable to suppose that, as we learn more about the world, we will learn more about how to learn. Granted the level 7 thesis, in other words, meta-knowability is a reasonable conjecture.

(ii) Meta-knowability is too good a possibility, from the standpoint of the growth of knowledge, not to be accepted initially, the idea only being reluctantly abandoned if all attempts at improving methods for the improvement of knowledge fail.

(i) and (ii) are not, perhaps, very strong grounds for accepting meta-knowability; both are open to criticism. But the crucial point, for the present argument, is that these grounds for accepting meta-

knowability, (i) and (ii), are independent of the success of science. This suffices to avoid circularity.²⁵³

But what about reasons for accepting theses at levels 5, 4 and 3? Are not these inevitably viciously circular? The thesis that the universe is comprehensible, at level 5, is accepted because no other idea, compatible with meta-knowability, has been so fruitful in generating empirically progressive research programmes; the thesis that the universe is physically comprehensible at level 4 is accepted because no other thesis, compatible with the level 5 thesis, has been so fruitful in generating empirically progressive research programmes;²⁵⁴ and likewise for the thesis at level 3. In short, theories at level 2 are accepted because of empirical success and compatibility with level 3, 4 or 5 theses; and these theses are

²⁵³. At this point I confess that in Maxwell (1998, p. 192-3) I give a third argument for accepting meta-knowability which does, perhaps, contain a whiff of circularity, in that it appeals to the *apparent* success of science. This suffices, I now think, to make this argument circular.

25. What justifies the claim that physicalism has been more fruitful for theoretical physics than any rival idea? This is justified by the point made in section II. All new, revolutionary, fundamental physical theories have been accepted because they (a) have brought greater unity to physics, and (b) have been more empirically successful, than any rivals – (a) being just as important as (b). In other words, *the* persisting non-empirical requirement for acceptance of revolutionary theory has been enhanced exemplification of physicalism (as far as theoretical physics as a whole is concerned) What irony that scientific revolutions – just that which convinced Kuhn (1970a) that there are ruptures in science with nothing theoretical surviving each rupture – actually demonstrate just the opposite: the persistent and increasingly successful search for unity, the assumption of underlying unity being repeatedly reinforced by each successive revolution. It may be asked: But how can revolutionary theories reinforce physicalism when the totality of physical theory has always, up till now, clashed with physicalism? The answer: If physicalism is true, then all physical theories that only unify a restricted range of phenomena, must be false. Granted the truth of physicalism, and granted that theoretical physics advances by putting forward theories of limited but ever increasing empirical scope, then it follows that physics will advance from one false theory to another, all theories being false until a unified theory of everything is achieved (which just might be true). The successful pursuit of physicalism requires progressive *increase* in both empirical scope and unity of the totality of fundamental physical theory. It is just this which the history of physics, from Galileo to today, exemplifies – thus demonstrating the unique fruitfulness of physicalism.

accepted because of their empirical fruitfulness. This would seem to be viciously circular in the most blatant fashion imaginable.

I have three arguments in refutation of this charge.

First, there is no question of the truth of theories being justified by an appeal to metaphysical theses, the truth of which is in turn justified by the success of science for the simple reason that AOE is thoroughly conjectural, and to that extent Popperian, in character, there being no attempt to justify the truth of either theories or metaphysical theses.

Secondly, physicalism is *incompatible* with accepted fundamental physical theories, so there could be no question of the truth of one being justified by an appeal to the truth of the other. Physicalism is deployed to *criticize*, and to try to *improve*, accepted fundamental theories, not to *justify* their truth.

Thirdly, and decisively, in so far as *acceptance* of physical theories is in part justified by an appeal to physicalism, whose acceptance is in turn justified by an appeal to the (apparent) success of science, which does involve a kind of circularity, this is licensed and legitimised by the level 6 thesis of meta-knowability. This asserts that the universe is such that there is some rationally discoverable thesis which, if accepted, makes possible the progressive improvement of more specific assumptions and methods in the light of the empirical success and failure of the research programmes to which they give rise. *If* meta-knowability is true, then progressively improving more specific metaphysical assumptions in the light of which seem to lead to the greatest empirical success, while at the same time choosing those empirically successful theories which best accord with these metaphysical assumptions, is just what needs to be done to make scientific progress. Meta-knowability, if true, justifies the element of circularity that is involved.

The gross invalidity of the genuinely viciously circular argument can be highlighted as follows. The argument seeks to justify acceptance of theory T by an appeal to metaphysical thesis M, and then justify acceptance of M by an appeal to the empirical success of T. But this argument works just as well (or ill) if we choose some empirically successful but horribly ad hoc rival to T, say T*, and a suitably ad hoc variant of M, say M*. We can now argue,

with equal validity (i.e. none): we justify acceptance of T* by appealing to M*, and justify acceptance of M* by appealing to the empirical success of T*. We have here a way of testing whether or not a putative solution to the problem of induction is, or is not, viciously circular: it must provide some valid way of ruling out arguments that appeal to ad hoc theories and theses like T* and M*.

AOE, granted the level 6 thesis of meta-knowability, does provide this. Given that M accords with meta-knowability in being rationally discoverable, all ad hoc rivals of M (i.e. M*) are ruled out because these are not “rationally discoverable”: M* is just one of infinitely many equally viable theses. Thus, *if* meta-knowability is accepted, AOE is not viciously circular – not circular in any invalid sense. Meta-knowability in effect asserts that the universe is such that no ad hoc or aberrant version of argumentation which appeals to T and M – a version which appeals to some T* and M* – can meet with success because all M*-type metaphysical theses are false.

It is of course absolutely vital that arguments for accepting meta-knowability do not themselves appeal to the success of science (for this would simply reintroduce vicious circularity at a higher level). The argument given above for accepting meta-knowability is weak, but it does not appeal, in any way whatsoever, to the success of science. Thus AOE is free of vicious circularity.²⁵⁵

IV

Even if AOE does not play van Fraassen’s Mug’s game, nevertheless how can it conceivably solve the *practical* problem of induction given its quasi-Popperian character?

Let me say at once that two versions of Popperianism deserve to be distinguished. On the one hand there is Popper's own view, which I shall call, with ironic intent, "dogmatic critical rationalism". This stresses merely the vital role that criticism has for rationality. Criticism is deployed, one might say, in an uncritical or almost dogmatic fashion. In contrast to this there is

²⁵⁵. For further details of the argument for AOE see Maxwell: (1998), (2001a, ch. 3 and app. 3), (2002d) and (2003d).

the version of critical rationalism which I wish to defend, which might be called "critical critical rationalism". This takes seriously the implications of a point emphasized, but not adequately followed up, by Popper, namely, that the whole point of *rational* criticism is to promote progress – and in connection with science, to promote progress in knowledge (and understanding). This means that theses which are demonstrably such that not accepting them can only harm and cannot help progress in knowledge whatever the universe is like, do not require (rational) criticism. They deserve to be permanently accepted. The cosmological thesis at level 7 is accepted on these grounds – in sharp contrast to anything found in Popper's work. Furthermore, it is all important, according to critical critical rationalism, to highlight that part of our knowledge which, we conjecture, it is most fruitful to criticize, from the standpoint of achieving progress in knowledge. Mere criticism is not good enough; we need to be critically critical, critical of criticism itself, directing criticism to that which we conjecture it is most fruitful to criticize from the standpoint of achieving progress. A basic idea behind the hierarchy of AOE is just to display the metaphysical presuppositions of science in such a way that that which, we conjecture, it is most fruitful to criticize be brought to the fore, fruitful criticism being especially facilitated. Criticism needs to be directed, above all, at theses at levels 1, 2 and 3 – theses from 4 to 6 becoming increasingly unfruitful to criticize as we ascend the hierarchy, due to their increasing lack of factual content and increasingly indispensable role in the search for knowledge.

V

But how does any of this help – the reader may ask with rising impatience – with solving the practical problem of induction? I now address this question head on.

It is important to appreciate that there are three parts to the problem of induction. There is the *methodological* problem: to specify the precise methods involved in the choice of theory in science. There is the *theoretical* problem: to show that we are justified in accepting the scientific theories we do accept, granted the aim is to improve theoretical knowledge and understanding of the universe. And, perhaps hardest of all, there is the *practical*

problem, the one just indicated above. The vast literature on the problem of induction is almost entirely devoted to the practical problem, but this is like trying to fly before you can crawl. The proper place to begin is with the methodological problem: if this has not been solved, to the extent of specifying *rigorous* methods for science, all efforts to solve the other two problems will be squandered on trying to justify the unjustifiable. This, in essence, is the reason for the long-standing failure of attempts to solve the problem(s) of induction.

Popper's falsificationism, like all other versions of SE, fails at the first hurdle (as I have, in effect, already pointed out above). Methods actually employed in physics involve persistently choosing unifying theories in preference to more empirically successful disunified rival theories. This in turn involves making a big, persistent metaphysical assumption, to the effect that all disunified theories are false. Rigour demands that this (implicit) metaphysical assumption be made explicit, within science, so that it can be criticized and, we may hope, improved. But falsificationism cannot do this, because its criterion of demarcation declares metaphysics to be non-scientific. And no version of SE can do this either, because the metaphysical assumption, implicit in persistent scientific preference for unifying theories against the evidence, is repudiated, denied, by all versions of SE. All SE views about scientific method lack rigour.

And related to the lack of rigour, there is a lack of precise characterization. Presuppositionism, as we have seen, leads to AOE, and to the view that metaphysical assumptions and associated methods *evolve* with evolving knowledge. No version of SE (including falsificationism) can do adequate justice to this evolving, positive feedback aspect of scientific method.²⁵⁶

²⁵⁶. Laudan (1977, 1984), inspired by Maxwell (1974) and earlier personal communication, does argue for changing methods of science within an SE view. But because of the anti-realist and SE character of his view, its lack of the hierarchical "meta-methodological" character of AOE, Laudan cannot do justice to the idea that new aims-and-methods need to be *appraised* so that those selected *improve* on earlier aims-and-methods, there being positive feedback between improving knowledge and improving knowledge-about -how-to-

Finally, in order merely to specify precisely what the methods of physics are it is necessary to specify what unification in theoretical physics *is*. We have seen in outline, above, what this amounts to: the more the totality of fundamental theory in physics exemplifies physicalism, so the more unified it is. What matters here is the *content* of physical theory, not its *form*. This solution to the problem of saying what theoretical unity *is* (which defeated even Einstein²⁵⁷) falls naturally out of AOE: but it is unavailable to SE (because it involves acknowledging that physicalism is a part of scientific knowledge). Notoriously, attempts to say what theoretical unity is, within the confines of SE, have failed dismally (see Salmon, 1989, for a review). SE, constitutionally incapable of characterizing theoretical unity, is incapable of specifying precisely those methods of physics which appeal to the "unity" of theory. Falsificationism, like all other versions of SE, for these three inter-related reasons, fails to solve even the *methodological* problem of induction – whereas AOE solves the problem with ease.

Having failed to solve the methodological problem it follows at once, of course, that falsificationism and other versions of SE all fail to solve the theoretical and practical problems of induction as well.

This has a bearing on the question of whether the problem of induction is solvable at all. Most contemporary philosophers, historians and sociologists of science seem to have concluded that the practical or justificational problem is unsolvable, just because repeated efforts to solve the problem seem to have got nowhere. But what the above point shows is that there is an entirely straightforward reason for this failure. All these failed attempts failed even to formulate the problem correctly! To do that one needs first to have identified the correct methods of science.

There is another way in which the problem may be misformulated so as to render it insoluble. The formulation may make epistemological demands that are so high that they are quite

improve-knowledge – a key feature of the rationality of science, according to AOE, and one which helps account for the explosive growth of modern science.

²⁵⁷. See Einstein (1969, pp. 20-25) and Maxwell (1998, pp. 105-6).

impossible to fulfill. One such formulation would be: How can our confidence that empirically successful, accepted scientific theories are true be justified? This makes impossibly high epistemological demands. All fundamental dynamical physical theories so far put forward, whatever empirical success they may have achieved, are false!²⁵⁸ Formulated in this way, the problem is insoluble. A slightly less epistemologically ambitious formulation would be: How can our confidence that the empirical predictions of empirically successful, accepted theories are true, be justified?” But this too asks for too much. All physical theories so far proposed, however empirically successful, yield false empirical predictions. A still less epistemologically ambitious formulation would be: How can our confidence that empirically successful, accepted theories yield true empirical predictions, within the standard range of phenomena (and accuracy) for which they have already been shown to be successful, be justified? But even this may ask for the impossible. Perhaps our customary confidence in science is misplaced. Perhaps just this is revealed by the correct solution to the practical or justificational problem.

In short, in order to avoid struggling to achieve the impossible, we need to formulate the problem in a somewhat more open-ended way than any of the above. The following stands a better chance of being solvable: How can our confidence that empirically successful, accepted theories yield true empirical predictions, within the standard range of phenomena (and accuracy) for which they have already been shown to be successful, be justified *in so far as such confidence is justified?*

29. Galileo's laws of terrestrial motion and Kepler's laws of planetary motion are contradicted by Newtonian theory, in turn contradicted by special and general relativity. The whole of classical physics is contradicted by quantum theory, in turn contradicted by quantum field theory. Science advances from one false theory to another. Viewed from an SE perspective, this seems discouraging and is often called "the pessimistic induction". Viewed from an AOE perspective, as I have already mentioned, this mode of advance is wholly encouraging, since it is required by AOE. Granted physicalism, the only way a dynamical theory can be precisely true of any restricted range of phenomena is to be true of all phenomena. All physical theories must be false until we obtain a theory of everything!

We cannot just assume, from the outset, that the solution to the problem must justify our “pre-Humean” confidence in common sense and scientific knowledge – the confidence we had, that is, before learning of Hume’s devastating arguments. For this may, again, be asking for the impossible. Perhaps Hume demonstrated, decisively, that such “pre-Humean” confidence is misplaced and unjustifiable. The correct solution to the justificational problem would, in this case, demonstrate just this to be the case. This, of course, is the view defended here.

We have seen that theoretical scientific knowledge makes assumptions about the cosmos. But there is more to it than that. Even very modest common sense knowledge, such as “I can walk across the room” or “This room will endure for the next ten seconds”, makes assumptions about the cosmos, about the nature of ultimate reality. Such very modest common sense theses imply (a) that the entire cosmos is such that no cosmic explosion is occurring anywhere which will spread with nearly infinite speed to engulf and destroy the earth in the next few seconds. We only possess knowledge of such modest common sense items in so far as we possess knowledge of the cosmological thesis (a). Science and common sense both make cosmological assumptions.

If these cosmological assumptions can be established to be true in such a way that we are justified in being confident of their truth, then there is the hope that we can be justifiably confident of the truth of both some empirical predictions of empirically successful theories, and modest items of common sense. But if the relevant cosmological theses are such that they cannot conceivably be established to be true in such a way that we are justified in being confident of their truth, then it follows straight away, from elementary logic, that we cannot conceivably be confident of the truth of either the predictions of scientific theories or the items of common sense. If our most modest, immediate, apparently secure items of common sense have implications for the nature of the entire cosmos that are irredeemably speculative and conjectural, then our modest items of common sense must themselves be irredeemably speculative and conjectural as well.

The relevant cosmological theses are indeed irredeemably speculative and conjectural, as everyone would surely admit. The

conclusion is thus inescapable: scientific knowledge, and modest common sense knowledge, are both irredeemably speculative and conjectural as well. If we demand of the solution to the justificational problem of induction that it restores our pre-Humean confidence in scientific and common sense knowledge, then we demand the impossible. The attempt to restore such pre-Humean confidence can only undermine the rationality of science, in so far as it lulls us into a false sense of security, and leads us to believe that parts of our knowledge do not need critical scrutiny.

One demand that can be made of the correct solution to the justificational problem of induction is that it puts empirical data and scientific theories onto an equal epistemological footing. Ordinarily we assume we can be confident of the truth of factual statements about our immediate, observed surroundings: this is a table, that is a book, this is a Bunsen burner, and so on. Before encountering Hume (and his 20th century descendents, such as Einstein, Popper and Kuhn), we may feel equally confident of the truth (or approximate truth) of empirically successful, accepted scientific theories. But Hume's argument, reformulated a little, has the effect of opening up a gulf between evidence and theory. Any theory has infinitely many empirical consequences – and consequences for distant times and places. However many consequences we verify empirically, we will always be infinitely far away from verifying the theory. Even if particular items of evidence are known with absolute certainty, theories must, it seems, remain pure conjectures. They can be falsified, perhaps, but remain permanently unverifiable, to any degree whatsoever. The solution to the justificational problem must, it may be felt, remove this gulf in epistemological status between empirical data and theory.

AOE does just that, by making clear that empirical data, just like theories, contain permanently conjectural cosmological implications. All knowledge, theoretical, empirical and common sense, is irredeemably conjectural because of these conjectural cosmological implications.

But if this is the case, what grounds are there for holding that AOE solves the justificational problem of induction? How does AOE do any better than Popper's conjecturalist position?

AOE, as have seen, solves the methodological problem, whereas Popper's falsificationism fails to do this. As a result, AOE is able to solve a very important part of the justificational problem which falsificationism, notoriously, fails to solve. This is the problem of discriminating decisively between those conjectures about whose truth we are so confident that we are prepared to entrust our lives to their truth, and conjectures about whose truth we have no confidence at all. Every time we fly in an aeroplane, cross a suspension bridge, or imbibe medicine, we entrust our lives to the correctness of relevant items of scientific knowledge, and confidently take for granted that rival conjectures that can be concocted (ostensibly even more empirically successful but horribly ad hoc) which predict we will die are all false. Accounting for this difference deserves to be regarded as the nub of the justificational problem. AOE accounts fully for this dramatic difference, whereas falsificationism entirely fails to do this.

John Worrall has dramatized the problem as follows. We are, let us suppose, standing on top of the Eiffel Tower, and we are confronted by two rival conjectures: one says if we jump we will float gently down to earth without harm; the other says we will fall in the usual way to our death (Worrall, 1989). Only lunatics think the first a viable possibility; the rest of us are absolutely confident in the truth of the second conjecture. How is this confidence to be justified? No version of SE comes up with an adequate answer, especially as aberrant versions of Newtonian theory or general relativity can be concocted which predict jumping on this occasion will lead to a soft, harmless landing, and which are empirically *more* successful than the non-aberrant versions of these theories. Can AOE justify our confidence that if we jump we will be killed?

If we grant the truth of the theses of AOE, from level 4 to 7, a straightforward answer can be given. Physicalism tells us that a unified pattern of physical law governs all phenomena. By far our best efforts at discovering invariant (or unified) laws governing such things as bodies in free fall near the earth's surface are Newton's theory of gravitation and, better still, Einstein's theory. No rival theory is even remotely as good at complying with the two requirements of (1) empirical success and (2) compatibility

with physicalism. Theories that are empirically more successful and predict a gentle landing can be concocted, but these clash horribly with physicalism, and deserve to be rejected for that reason. But Newton's or Einstein's theory (plus additional information about such things as the mass of the earth) predict with stark clarity: jumping leads to rapid acceleration at roughly 32 ft per sec². Above all, a theory which accords with physicalism as well as Newton's or Einstein's theory, but predicts that jumping will lead to a gentle floating to the ground is nowhere on the horizon. Thus, given the truth of physicalism, there is absolutely no question, no grounds for serious doubt, whatsoever: jumping is for suicides only.

But we are not given the truth of physicalism. At most arguments deployed above give grounds for accepting physicalism granted our aim is to improve our conjectural knowledge of truth. There are arguments justifying *acceptance* of theses at levels 3 to 7, but no arguments justifying the *truth* of these theses. And it is the latter we require, it seems, to solve Worrall's problem, and the practical problem of induction more generally.

I have two replies to this objection.

First, even in the absence of any kind of justification of the truth of physicalism AOE succeeds, nevertheless, in distinguishing decisively between conjectures we are confident are true, to the extent even of entrusting our lives to their truth, and conjectures (even empirically more successful conjectures) about whose truth we have no such confidence.

Second, the demand that the truth of physicalism must receive some kind of justification before it becomes rationally acceptable for practical purposes is not just impossible to fulfil; it deserves to be rejected in that it stems from an unrigorous, untenable conception of science, and human knowledge more generally. If, and only if, some version of SE is correct, and science is based on evidence, and on metaphysical theses whose truth has been justified (if there are any), is the demand to justify the truth of physicalism itself justifiable. But all versions of SE are unrigorous and untenable. Hence, the SE demand to justify physicalism is itself unjustifiable, and must be rejected. What has been

demonstrated above is that *all* significant factual knowledge, common sense and scientific, implies (and thus presupposes) cosmological theses: rigour requires that these unjustifiable cosmological theses are made explicit, so that they can be critically assessed and, we may hope, improved. To demand that such cosmological theses cannot be accepted unless their truth is justified condemns science to lack of rigour, because it ensures that unjustifiable cosmological theses will not be, and cannot be, accepted as a part of scientific knowledge. The demand deserves to be rejected.

Human knowledge has always had this inescapable cosmological dimension built into it. The illusion that science could dispense with such unjustifiable cosmological conjectures only crept in with the general acceptance of SE, some time after Newton and before the end of the 19th century. What needs to be done is not to justify the truth of physicalism, but rather to justify the claim that this cosmological conjecture has played a more fruitful role in the advance of science than any rival at that level. Just this was done above. *Science does not prove its cosmological conjectures; it sets out to improve those it has inherited from the past.* Physicalism is the best available, at that level of generality, and that suffices to solve the Worrall problem, and the justificational problem of induction. We are justified in entrusting our lives to the standard empirical predictions of those theories (a) which have met with sufficient empirical success, and (b) which, together with other such empirically successful theories, are more nearly compatible with our best metaphysical theses concerning the comprehensibility and knowability of the universe. Our best metaphysical theses, in turn, are those which have generated the most empirically progressive scientific research programmes. The circularity that seems to be involved here is legitimised by acceptance of meta-knowability.

VI

I now spell out in a little more detail the point just made, central to the solution to the practical problem of induction.

When viewed from the perspective of SE, it looks as if, in order to solve the practical problem, sufficiently good grounds must be given for the *truth* of physicalism, from scratch as it were.

Acknowledging, initially, only knowledge of particular empirical facts, we must somehow provide an argument for the truth of physicalism which is so good that it justifies us in being entirely confident of the truth of physicalism even when our lives are at stake. Let us call this the "SE requirement" for solving the practical problem of induction. Given this requirement, the prospects for solving the practical problem of induction seem hopeless.

But this requirement itself deserves to be rejected. It is only acceptable if what I shall call "the SE prescription" is acceptable. But this latter is an intellectual disaster and deserves to be rejected. Hence the SE requirement must be rejected as well.

By the SE prescription I mean this: in order to develop science in a properly scientific, rigorous way, so that it is capable of delivering authentic, reliable knowledge, science must eschew all metaphysical presuppositions in the context of justification and base acceptance of scientific theories as far as possible solely on empirical considerations without reference to conjectural metaphysics. If this is correct, then it makes perfect sense to demand of any attempt to solve the practical problem of induction that it satisfies the SE requirement. But the SE prescription – eschewing metaphysical presuppositions in order to render science scientific and rigorous – is, we have seen, an intellectual disaster. It has entirely the opposite effect of the one intended. If taken seriously, instead of enhancing the rigour of science, it would destroy science and, indeed, all knowledge. The arguments of this and previous chapters demonstrate that science devoid of metaphysics is not possible. Selecting theories on the basis of empirical success and failure, no kind of assumption being made about what kind of world this is, cannot succeed if rigorously pursued because science would be overwhelmed by endlessly many empirically successful aberrant theories which would stultify science and render technological application impossible. (Or, if requirements of simplicity and unity are invoked, in addition to empirical requirements, then metaphysical assumptions of unity are being presupposed, but in a surreptitious fashion.) Science has survived and progressed despite, and not because of, acceptance of SE by the scientific community. Science has managed to do this by

implementing SE in only a highly unrigorous and hypocritical fashion (implicit metaphysical presuppositions exercising a highly influential role over choice of theory).

The SE prescription is, then, an intellectual disaster. Unfounded metaphysical or cosmological conjectures about the comprehensibility and knowability of the universe are essential for science, and cannot be eliminated without disaster. What needs to be implemented, instead, is the "AOE prescription": roughly, endeavour to *improve* metaphysical assumptions explicit or implicit in current science and knowledge; do this by modifying existing assumptions in that direction which seems to be the most fruitful from the standpoint of acquiring empirical knowledge within a fixed framework of assumptions and methods which are such that these are required for the acquisition of knowledge to be possible at all. In other words, put generalized AOE into practice.

Before the advent of SE, the pursuit of science, or of knowledge more generally, invariably went on within a framework of religious and metaphysical assumptions. Christianity, the corpuscular hypothesis, and Galileo's thesis that the book of nature is written in the language of mathematics, played an especially important role in the sixteenth and seventeenth centuries as far as the birth and development of modern science was concerned. The orthodox SE prescription insists that science must be dissociated from such dubious, unfounded religious and metaphysical doctrines. But this cannot be done; at best, unrigorous and hypocritical science results. The impression that it has been done creates the insoluble problem of induction (as traditionally construed, from the SE perspective).

Instead, we need to see AOE science as explicitly *improving* on antecedently upheld religious and metaphysical theses. There is, in science, a substantial component of faith – but, ideally, it is *rational* faith, openly acknowledged as conjectural in character, subjected to sustained criticism, and undergoing persistent modification in the direction of that which seems to lead to the most empirically progressive research, at levels 1 and 2. Science does not eliminate metaphysics; it implements a method which makes it possible for us to develop and choose those metaphysical ideas most fruitful for progress in empirical knowledge.

But once the SE prescription has been rejected as intellectually disastrous, and the AOE prescription accepted instead, it is clear that the SE requirement for solving the practical problem of induction must itself be rejected as intellectually unreasonable and unacceptable. The only rationale for adopting it arises from the idea that it is entirely proper to put the SE prescription into practice. But putting this into practice makes science impossible (for reasons wholly in addition to the resulting insolubility of the practical problem of induction). The SE requirement presumes a state of knowledge that has resulted from implementing the intellectually destructive SE prescription: this state of knowledge is an intellectual disaster, and must be rejected, and along with it the SE requirement.

Taking the SE requirement seriously is rather like an athlete having both legs amputated and then expecting to win the 100 metres at the Olympics. Render science, and indeed all knowledge, impossible (by throwing away vital metaphysics) and it should occasion no surprise that a situation is created in which the practical problem of induction becomes impossible to solve as well.

For the last four centuries ever since Galileo, or perhaps for the last two and a half thousand years ever since the Presocratics, physicalism has been by far the most fruitful metaphysical thesis available from the standpoint of promoting progress in science, or in knowledge more generally. No rival metaphysical thesis has been remotely as fruitful. AOE correctly depicts our current scientific knowledge. Physicalism, in short, is justifiably a basic tenet of current (conjectural) scientific knowledge, our best attempt, at that vital level, of acquiring knowledge of truth, more secure, indeed, than any fundamental physical theory, such as quantum theory or general relativity. It makes very good sense not to jump off the Eiffel tower if you want to stay alive, for the reasons given above.

It may be asked: but if science must accept, and not eschew, religious or metaphysical ideas, what grounds are there for preferring physicalism to a thesis such as that God exists and benevolently arranges for the cosmos to be such as to make science

possible? For my reply see Maxwell (1998, pp. 206-8; 2001a, pp. 6-10; 2002c).

The argument given above is a special case of a more general argument concerning Cartesianism. The proper fundamental problem of epistemology is the problem of how we can best *improve* knowledge (improve what we have inherited from the past). The Cartesian prescription says: throw everything away we have inherited from the past except that which cannot be doubted, or is most secure: taking this as a secure base, build up rigorous, reliable knowledge. This Cartesian prescription has exercised a profound influence on philosophy, on both so-called rationalists and on empiricists. Influencing the former it leads to the search for secure principles founded on reason; influencing the latter it leads to the idea that observational knowledge alone constitutes the only acceptable Cartesian secure base, it being necessary to build the rest of knowledge up from this secure base. This, of course, is the SE prescription; it is a version of the Cartesian prescription. But the Cartesian prescription must be rejected. The proper way to set about *improving* knowledge is to acknowledge the conjectural character of what we have inherited from the past, subject it to critical scrutiny, and put generalized AOE into practice. In direct opposition to Cartesianism, this involves in part taking most seriously ideas that are most *vulnerable* to being found to be false, namely the falsifiable theories of science. Far from giving priority to ideas immune to doubt, we need to give priority to ideas most vulnerable to doubt. In this way, as Popper stresses, we make it possible for us to learn from our mistakes. But this Popperian prescription needs itself to be modified, as we have seen, so that some unfalsifiable, metaphysical ideas, inherited from the past, continue to be accepted and developed as a part of scientific knowledge, those in particular being accepted which either (a) must be true if knowledge is to be possible at all, or (b) are most fruitful in leading to progress in empirical knowledge. The Cartesian prescription deserves to be resoundingly rejected, and along with it the SE prescription and SE requirement.

The problem of induction is not just a philosophical puzzle à la Wittgenstein. Its long-standing insolubility is indicative of a fundamental defect in our understanding of science and its

relationship with metaphysics and philosophy – a fundamental defect in our whole culture. Science does not stand opposed to metaphysics and philosophy; it is metaphysics and philosophy carried on employing the improved methods of investigation of empiricism: observation and, above all, experimentation (a point enshrined in the 17th century terms of "experimental" and "natural" philosophy). A basic task for philosophers today is to try to get across to the scientific community just how vital metaphysics and philosophy, properly conducted, are for science, so that scientists and philosophers can begin to collaborate on implementing AOE science, thus recreating natural philosophy. But this is unlikely to happen as long as AOE continues to be dismissed, so unjustly, as a mug's game.

Chapter Thirteen

Does Probabilism Solve the Great Quantum Mystery?

(First published in *Theoria* vol. 19/3, no. 51, 2004, pp. 321-336.)

Abstract

What sort of entities are electrons, photons and atoms given their wave-like and particle-like properties? Is nature fundamentally deterministic or probabilistic? Orthodox quantum theory (OQT) evades answering these two basic questions by being a theory about the results of performing measurements on quantum systems. But this evasion results in OQT being a seriously defective theory. A rival, somewhat ignored strategy is to conjecture that the quantum domain is fundamentally probabilistic. This means quantum entities, interacting with one another probabilistically, must differ radically from the entities of deterministic classical physics, the classical wave or particle. It becomes possible to conceive of quantum entities as a new kind of fundamentally probabilistic entity, the “propensiton”, neither wave nor particle. A fully micro realistic, testable rival to OQT results.

1 Orthodox Quantum Theory is the Best and Worst of Theories

What sort of entities are electrons, photons and atoms given their wave-like and particle-like properties? Is nature fundamentally deterministic or probabilistic? Any decent theory of the quantum domain, able to provide us with genuine knowledge and understanding of its nature, ought to provide answers to these childish elementary questions. Orthodox quantum theory (OQT) evades answering these questions by being a theory, not about

quantum systems as such, but rather about the results of performing measurements on such systems.²⁵⁹

This state of affairs came about as follows. Bohr, Heisenberg, Dirac and the other creators of OQT did not know how to solve the quantum wave/particle dilemma. This created a grave problem for those seeking to develop quantum theory. How can one develop a consistent theory about entities that seem to be both wave-like and particle-like, as in the two slit experiment for example? Heisenberg around 1925 hit upon the strategy of evading this fundamental dilemma by developing what subsequently became matrix mechanics as a theory exclusively about the results of performing measurements on quantum systems, this version of quantum theory thus not needing to specify the nature of quantum systems when not undergoing measurement. Schrödinger, a little later in 1926, developed wave mechanics in the hope that it would be a precise theory about the nature of quantum systems. This theory, Schrödinger hoped, would show the electron to be wave-like in character. But then Born successfully interpreted the Ψ function of Schrödinger's wave mechanics as specifying the probability of detecting the particle in question. According to Born's crucial interpretative postulate, $|\Psi|^2 \cdot dV$ gives the probability of detecting the particle in volume element dV if a position measurement is performed. Schrödinger proved that his theory and Heisenberg's matrix mechanics are equivalent: the outcome, a sort of synthesis of the two theories, is OQT.

OQT is an extraordinarily successful theory empirically, perhaps the most successful in the whole of physics when one takes into account the range, immense diversity, and accuracy of its predictions. But not only does it fail to solve the great quantum mystery of what sort of entities electrons and atoms can be in view of their apparently contradictory particle and wave properties. It also fails to answer the other childish elementary question: Is the quantum domain deterministic or probabilistic? The basic

1. Good introductory accounts of OQT, increasingly technical, are Squires (1986); Gillespie (1973); Feynman et al (1965). See also Maxwell (1998, appendix).

dynamic equation of OQT, Schrödinger's time-dependent equation, is deterministic in character. It tells us that quantum states, given by Ψ , evolve deterministically in time, as long as no measurements are made. But this does not mean OQT asserts that the quantum domain is deterministic. First, given OQT, Ψ cannot be interpreted as specifying the actual physical state of a quantum system, just because OQT fails to solve the wave/particle dilemma, and thus fails to provide a consistent specification of the physical nature of quantum systems when not being measured. Given OQT, Ψ must be interpreted as containing no more than information about the outcome of performing measurements. Secondly, OQT in general makes *probabilistic* predictions about the outcome of performing measurements, not (apart from exceptional circumstances) *deterministic* predictions. But one cannot conclude from this that OQT asserts that the quantum domain is fundamentally *probabilistic* in character, some physical states of affairs only determining what occurs subsequently only *probabilistically*. This is because, according to OQT, probabilistic outcomes only occur when we intervene, and make a *measurement*. In the absence of measurement, nothing probabilistic occurs at all, according to OQT. Indeed, if the process of measurement is treated quantum mechanically, then nothing probabilistic occurs at all, precisely because the basic dynamic equation of OQT, Schrödinger's time-dependent equation, is deterministic.

The inability of OQT to answer these two elementary questions is in itself a serious failure of the theory. But there are, as a consequence, a host of further failures and defects. Because OQT is about the results of performing measurements on quantum systems (and not about quantum systems *per se*, due to its failure to solve the wave/particle problem), in order to come up with physical predictions OQT must consist of two parts, (1) quantum postulates, and (2) some part of classical physics for a treatment of measurement. (2) is indispensable. (1) alone, precisely because OQT lacks its own quantum ontology, cannot predict anything physical at all – or at least can only make conditional predictions of the form: if such and such a measurement is made, such and such will be the outcome with such and such a probability. Thus

OQT = QP + CP, where “QP” stands for the quantum mechanical postulates of the theory, and “CP” stands for the classical postulates, required for measurement.

In what follows, a quantum “measurement” is a process that actually detects quantum systems; a process which prepares a quantum system to be in a certain quantum state, but does not detect the system, is a “preparation” rather than a “measurement”.

OQT, construed as QP + CP, as it must be, is a seriously defective theory. (a) OQT is *imprecise*, due to the inherent lack of precision of the notion of “measurement”. How complex and macroscopic must a process be before it becomes a measurement? Does the dissociation of one molecule amount to a measurement? Or must a thousand or a million molecules be dissociated before a measurement has been made? Or must a human being observe the result? No precise answer is forthcoming. (b) OQT is ambiguous, in that if the measuring process is treated as a measurement, the outcome is in general probabilistic, but if this process is treated quantum mechanically, the outcome is deterministic. (c) OQT is very seriously *ad hoc*, in that it consists of two incompatible, conceptually clashing parts, QP and CP. OQT only avoids being a straightforward contradiction by specifying, in an arbitrary, *ad hoc* way, that QP applies to the quantum system up to the moment of measurement, and CP applies to the final measurement result. (d) OQT is non-explanatory, in part because it is *ad hoc*, and no *ad hoc* theory is fully explanatory, in part because OQT must presuppose some part of what it should explain, namely classical physics. OQT cannot fully explain how classical phenomena emerge from quantum phenomena because some part of classical physics must be presupposed for measurement. (e) OQT is limited in scope in that it cannot, strictly speaking, be applied to the early universe in conditions which lacked preparation and measurement devices. Strictly speaking, indeed, it can only be applied if physicists are around to make measurements. (f) OQT is limited in scope in that it cannot be applied to the cosmos as a whole, since this would require preparation and measurement devices that are outside the cosmos, which is difficult to arrange. Quantum cosmology, employing OQT, is not possible. (g) For somewhat similar reasons, OQT is such that it resists unification with general

relativity. Such a unification would presumably involve attributing some kind of quantum state to spacetime itself (general relativity being a theory of spacetime). But, granted the basic structure of OQT, this would require that preparation and measurement devices exist outside spacetime, again not easy to arrange.

These nine defects, the two basic failures with which we began and the seven consequential defects, (a) to (g), are, taken together, very serious indeed. Despite its immense empirical success, OQT must be declared to be an unacceptably defective theory. It is the best of theories, and the worst of theories.²⁶⁰

In opposition to this conclusion, it may be argued that all physical theories, even a classical theory such as Newtonian theory (NT), must call upon additional theory to be tested empirically. In testing predictions of NT concerning the position of a planet at such and such a time, optical theory is required to predict the results of telescopic observations made here on earth. But this objection misses the point. NT is perfectly capable of issuing in *physical* predictions without calling upon additional theory, just because it has its own physical ontology. NT, plus initial and boundary conditions formulated in terms of the theory, can issue in the *physical* prediction that such and such a planet is at such and such a place at such and such a time, whether anyone observes the planet or not, without calling upon optical theory or any other theory. This OQT cannot do. It cannot do this because it lacks its own quantum ontology, having failed to solve the quantum wave/particle problem. In order to deliver an unconditional physical prediction, OQT must call upon some part of classical physics, as a matter of necessity, so that the theory can refer to something physically actual. The case of NT and OQT are quite different, because NT postulates actually existing physical bodies

2. Rival interpretations of quantum theory include: Bohm's interpretation, according to which quantum systems are both particles and waves; Everett's many-worlds interpretation; decoherence; consistent histories. None of these, in my view, provides us with a satisfactory version of quantum theory. For critical surveys and further literature see Squires (1986); Rae (2002, ch. 13); Bacciagaluppi (2003).

whether observed or not, whereas QP does not; for that one requires OQT, that is QP + CP.

It may be objected that even if non-relativistic quantum theory fails to solve the wave/particle problem, relativistic quantum theory, or quantum field theory, does solve the problem in that it declares that what exists is the quantum field, “particles” being discrete excitations of the field. But this objection misses the point as well. Orthodox quantum field theory (OQFT) is just as dependent on measurement, and thus on some part of classical physics, as non-relativistic OQT is. The quantum states of the quantum field of OQFT have to be interpreted as making probabilistic predictions about the results of performing measurements, just as in the case of OQT. A version of quantum field theory which succeeded in specifying the nature of the quantum field in a fully satisfactory way, so that the theory has its own quantum ontology entirely independent of any part of classical physics, would be able to issue in physical predictions about actual physical states of affairs entirely independently of measurement. Such a theory would be able to predict and explain macroscopic, quasi-classical phenomena as arising from the quantum field alone, without calling upon some part of classical physics for a treatment of measurement. This OQFT cannot do.

2 Probabilism to the Rescue

What needs to be done to cure OQT of its serious defects? The primary task must be to specify precisely and unambiguously the nature of quantum entities so that quantum theory (QT) can be formulated as a testable theory about how these entities evolve and interact without there being any mention of measurement or observables in the postulates of the theory at all. The key point that needs to be appreciated, I suggest, in order successfully to complete this task, is that the quantum domain is fundamentally probabilistic.²⁶¹ It is this that the manifestly probabilistic character of QT is trying to tell us.

3. Popper has suggested that probabilism is the key to understanding wave/particle duality, and has put forward a propensity interpretation of quantum theory: see Popper (1957b,

The approach to solving the mysteries of the quantum domain that I am suggesting here has been long ignored largely because of the accidents of history. When Quantum Theory (QT) was being developed and interpreted, during the first three decades of the last century, two opposing camps developed: the Bohr-Heisenberg camp, which argued for the abandonment of micro-realism, and the abandonment of determinism; and the Einstein-Schrödinger camp, which argued for the retention of realism, and the retention of determinism. One result of this polarization of views was that the idea of *retaining* realism but *abandoning* probabilism got overlooked. But it is just this overlooked option, I maintain, which gives us our best hope of curing the defects of QT. One might call this option *probabilistic micro-realism*.

Once we acknowledge that the quantum domain is fundamentally probabilistic, so that the basic laws governing the way quantum systems interact with one another are probabilistic, it is clear that *measurement* cannot be a satisfactory necessary and sufficient condition for probabilistic transitions to occur. Probabilistic transitions must be occurring in nature whether or not physicists are around to dub certain processes “measurements”. The very notion of measurement is in any case, as we have seen, inherently imprecise. We require a new, precise, necessary and sufficient condition for probabilistic transitions to occur, to be specified in fundamental, quantum mechanical terms.

Furthermore, once the fundamentally probabilistic character of the quantum domain is acknowledged, it immediately becomes clear how the key quantum wave/particle problem is to be solved. If the quantum domain is fundamentally probabilistic, then the physical entities of this domain, electrons, atoms and the rest, cannot possibly be classical, deterministic entities – classical particles, waves or fields. Quite generally, we should hold that there is a one-to-one correspondence between the dynamical laws of a physical theory on the one hand, and the entities and their

1967, 1982). His interpretation of quantum theory is, however, unsatisfactory and quite different from the one I advocate here. For my criticisms of Popper see Maxwell (1976b, 285-6; 1985, 41-2).

physical properties postulated by the theory, on the other hand. In speaking of the entities, and the properties of entities, postulated by a physical theory, we are thereby speaking, in other terms, of the dynamical laws of the theory. Hence, change dynamical laws in some basic way, and we thereby change postulated physical entities and their properties. In particular, change dynamical laws dramatically, so that they become probabilistic instead of being deterministic, and the nature of postulated physical entities must change dramatically as well. Quantum entities, interacting with one another probabilistically, must be quite different from all physical entities so far encountered within deterministic classical physics.

[Elsewhere (Maxwell, 1976b, 283-6; 1988, 44-8) I have indicated how the notion of probabilistic physical property, or propensity, that is being presupposed here, amounts to a probabilistic generalization of the notion of deterministic, necessitating property explicated in Maxwell (1968); see also Maxwell (1998, 141-55). I might add, no doubt controversially, that in my view my 1968 paper gives the definitive account of how dispositional, necessitating properties in physics should be conceived. This viewpoint, in particular, makes no appeal to Kripke's (1981) fallacious considerations concerning identity and necessity: for a refutation of Kripke, see Maxwell (2001a, appendix 2). Much subsequent work on dispositional properties in science is vitiated by a failure to take my earlier work into account, and a reliance instead on Kripke.]

The defects of OQT have arisen, in other words, because physicists have sought to interpret probabilistic quantum theory in terms of classical waves and particles, deterministic metaphysical ideas appropriate to earlier classical physics but wholly inappropriate to the new quantum theory. The failure of this entirely misguided attempt then led to despair at the possibility of solving the (misconstrued) wave/particle problem, despair at the possibility of specifying the precise physical nature of quantum entities. This despair in turn led to the development of OQT as a theory about the results of performing measurements – a theory which, it seemed, did not need to specify the precise nature of

quantum entities. But the outcome is a theory burdened with the nine serious defects indicated above.

Thus the traditional quantum wave/particle problem is the *wrong* problem to pose. We should ask, not “Are quantum entities waves or particles?”, but rather (1) What kinds of possible, unproblematic, fundamentally probabilistic physical entities are there?”, and (2) Are quantum entities one kind of such unproblematic probabilistic entity?

The failure to put right the serious defects of QOT has persisted for so long because physicists have abandoned hope of solving the traditional quantum wave/particle problem, not realizing that this is entirely the *wrong* problem to try to solve in the first place. Once it is appreciated that (1) and (2) are the *right* problems to try to solve, new possibilities, long overlooked, immediately spring to mind.

First, physical entities that interact with one another probabilistically may be dubbed *propensitons*. Two kinds of unproblematic propensiton can immediately be distinguished: *continuous* propensitons, which evolve probabilistically continuously in time, and *intermittent* propensitons, which evolve deterministically except for intermittent moments in time when appropriate physical conditions arise, and the propensitons undergo probabilistic transitions.

There is a second obvious distinction that can be made between propensitons which spread out spatially in time, increasing the volume of space they occupy with the passage of time, and propensitons which do not spread spatially in this way. Let us call the first *spatially spreading* propensitons, and the second *spatially confined* propensitons.

We are in new territory. In our ordinary experience of the world, and within deterministic physics, we never encounter propensitons. Probabilistic outcomes, obtained when we toss a penny or a die, can always be put down to probabilistic changes in initial conditions. Classical statistical mechanics presupposes that the underlying dynamic laws are deterministic. Having no experience of them, propensitons will, inevitably, when we first encounter them, strike us as mysterious, even unacceptably weird. But these feelings of unfamiliarity ought not to lead us into

deciding that theories which postulate such entities are inherently unacceptable. In particular, the four kinds of propensity indicated above should be regarded as equally viable, *a priori*. Whether a theory that postulates one or other type of propensiton is acceptable or not should be decided upon in the usual way, in terms of its empirical success, and the extent to which it is unified, simple, explanatory.

Granted that quantum systems are *some* kind of propensiton, which of the four kinds of unproblematic propensiton just indicated should we take quantum systems to be? There is here a very important consideration to be borne in mind. Despite suffering from the nine defects indicated above, nevertheless OQT is perhaps the most empirically successful physical theory ever formulated. The range, variety and accuracy of its empirical predictions are unprecedented. No other physical theory has been subjected to such sustained severe experimental testing, and has survived without a single refutation. There are good grounds for holding that OQT has got quite a lot right about the nature of the quantum world. Our strategy, then, ought to be, in the first instance at least, to stick as close to OQT as possible, and modify OQT just sufficiently to remove the defects of the theory. The structure of OQT mirrors that of the intermittent, spatially spreading propensiton. On the one hand quantum states evolve deterministically, in accordance with Schrödinger's time-dependent equation; on the other hand, there are, on the face of it, probabilistic transitions associated with measurement. Quantum states spread out spatially when evolving deterministically, and tend to become localized when measurements are made. All this mirrors the character of the intermittent, spatially spreading propensiton, the only unsatisfactory feature of OQT being that the theory stipulates that probabilistic transitions occur when measurements are made.

A very elementary kind of spatially spreading intermittent propensiton is the following. It consists of a sphere, which expands at a steady rate (deterministic evolution) until it touches a second sphere, at which moment the sphere becomes instantaneously a minute sphere, of definite radius, somewhere within the space occupied by the large sphere, probabilistically

determined. The second sphere undergoes the same instantaneous probabilistic transition. Then both minute spheres again undergo steady, deterministic expansion, until they touch once more, and another probabilistic localization occurs.

A slightly more sophisticated version of this elementary spatially spreading intermittent propensiton is the following. The sphere is made up of variable “position probability density”, such that, when the sphere localizes probabilistically, in the way just indicated, it is most probable that it will be localized where the position probability density is most dense. A law specifies how position probability density is distributed throughout the sphere. We might even imagine that the position probability density exhibits a wave-like distribution. Such a propensiton, given appropriate conditions for probabilistic localization, might even exhibit interference phenomena in a two-slit experiment!

Quantum entities, such as electrons, photons and atoms, are, I suggest, spatially spreading intermittent propensitons. Their physical state is specified by the Ψ function of QT. The deterministic evolution of these quantum propensitons is specified by Schrödinger’s time-dependent equation:-

$$i\hbar \frac{\partial \Phi(t)}{\partial t} = - \frac{\hbar^2}{2m} \nabla^2 \Phi(t) + V \Phi(t).$$

The crucial questions that need to be answered to specify precisely the probabilistic properties – or propensities – of quantum systems are these:

- (a) What is the precise quantum mechanical condition for a probabilistic transition to occur?
- (b) Given the quantum state, Ψ , at the instant before the probabilistic transition, how does this determine what the possible outcome states are, $\phi_1, \phi_2, \dots \phi_N$?
- (c) How does Ψ determine the probability p_r that the outcome of the probabilistic transition will be ϕ_r , for $r = 1, 2, \dots N$?

(d) How can (a) to (c) be answered so that the resulting fundamentally probabilistic version of quantum theory reproduces all the empirical success of OQT?

A number of different answers can be given to (a) to (d).

One possibility is the proposal of Ghirardi, Rimini and Weber (see Ghirardi and Rimini, 1990) according to which the quantum state of a system such as an electron

collapses spontaneously, on average after the passage of millions of years, into a highly localized state. When a measurement is performed on the quantum system, it becomes quantum entangled with millions upon millions of quantum systems that go to make up the measuring apparatus. In a very short time there is a high probability that one of these quantum systems will spontaneously collapse, causing all the other quantum entangled systems, including the electron, to collapse as well. At the micro level, it is almost impossible to detect collapse, but at the macro level, associated with measurement, collapse occurs very rapidly all the time.

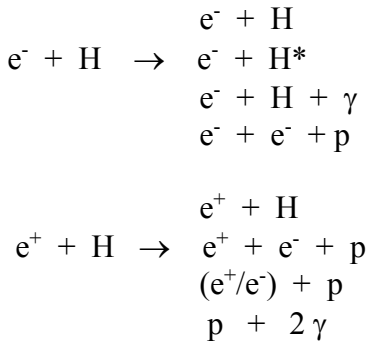
Another possibility is the proposal of Penrose (1986), according to which collapse occurs when the state of a system evolves into a superposition of two or more states, each state having, associated with it, a sufficiently large mass located at a distinct region of space. The idea is that general relativity imposes a restriction on the extent to which such superpositions can develop, in that it does not permit such superpositions to evolve to such an extent that each state of the superposition has a substantially distinct space-time curvature associated with it.

The possibility that I favour, put forward before either Ghirardi, Rimini and Weber's proposal, or Penrose's proposal, is that probabilistic transitions occur whenever, as a result of inelastic interactions between quantum systems, new "particles", new bound or stationary systems, are created (Maxwell, 1972b, 1976b, 1982, 1988, 1994). A little more precisely:

Postulate 1A: Whenever, as a result of an inelastic interaction, a system of interacting "particles" creates new "particles", bound or stationary systems, so that the state of the system goes into a

superposition of states, each state having associated with it different particles or bound or stationary systems, then, when the interaction is nearly at an end, spontaneously and probabilistically, entirely in the absence of measurement, the superposition collapses into one or other state.

Two examples of the kind of interactions that are involved here are the following:



(Here e^- , e^+ , H , H^* , γ , p and (e^+/e^-) stand for electron, positron, hydrogen atom, excited hydrogen atom, photon, proton and bound system of electron and positron, respectively.)

What exactly does it mean to say that the “interaction is very nearly at an end” in the above postulate? My suggestion, here, is that it means that forces between the “particles” are very nearly zero, except for forces holding bound systems together. In order to indicate how this can be formulated precisely, consider the toy interaction:

$$\begin{array}{l}
 a + b + c \rightarrow a + b + c \quad (A) \\
 a + (bc) \quad (B)
 \end{array}$$

Here, a , b and c are spinless particles, and (bc) is the bound system. Let the state of the entire system be $\Phi(t)$, and let the asymptotic states of the two channels (A) and (B) be $\psi_A(t)$ and $\psi_B(t)$ respectively. Asymptotic states associated with inelastic interactions are fictional states towards which, according to OQT,

the real state of the system evolves as $t \rightarrow +\infty$. Each outcome channel has its associated asymptotic state, which evolves as if forces between particles are zero, except where forces hold bound systems together.

According to OQT, in connection with the toy interaction above, there are states $\phi_A(t)$ and $\phi_B(t)$ such that:

- (1) For all t , $\Phi(t) = c_A\phi_A(t) + c_B\phi_B(t)$, with $|c_A|^2 + |c_B|^2 = 1$;
- (2) as $t \rightarrow +\infty$, $\phi_A(t) \rightarrow \psi_A(t)$ and $\phi_B(t) \rightarrow \psi_B(t)$.

The idea is that at the first instant t for which $\phi_A(t)$ is very nearly the same as the

asymptotic state $\psi_A(t)$, and $\phi_B(t)$ is very nearly the same as $\psi_B(t)$, then the state of the system, $\Phi(t)$, collapses spontaneously either into $\phi_A(t)$ with probability $|c_A|^2$, or into $\phi_B(t)$ with probability $|c_B|^2$. Or, more precisely:

Postulate 1B: At the first instant for which $|\langle \psi_A(t) | \phi_A(t) \rangle|^2 > 1 - \epsilon$

or $|\langle \psi_B(t) | \phi_B(t) \rangle|^2 > 1 - \epsilon$, the state of the system collapses spontaneously into $\phi_A(t)$ with probability $|c_A|^2$, or into $\phi_B(t)$ with probability $|c_B|^2$, ϵ being a universal constant, a positive real number very nearly equal to zero.

The evolutions of the actual state of the system, $\Phi(t)$, and the asymptotic states, $\psi_A(t)$ and $\psi_B(t)$, are governed by the respective channel Hamiltonians, H , H_A and H_B , where:-

$$H = - \left(\frac{\hbar^2}{2m_a} \nabla_a^2 + \frac{\hbar^2}{2m_b} \nabla_b^2 + \frac{\hbar^2}{2m_c} \nabla_c^2 \right) + V_{ab} + V_{ac} + V_{bc}$$

$$H_A = - \left(\frac{\hbar^2}{2m_a} \nabla_a^2 + \frac{\hbar^2}{2m_b} \nabla_b^2 + \frac{\hbar^2}{2m_c} \nabla_c^2 \right)$$

$$H_B = - \left(\frac{\hbar^2}{2m_a} \nabla_a^2 + \frac{\hbar^2}{2m_b} \nabla_b^2 + \frac{\hbar^2}{2m_c} \nabla_c^2 \right) + V_{bc}$$

Here, m_a , m_b , and m_c are the masses of “particles” a, b and c respectively, and $\hbar = h/2\pi$ where h is Planck’s constant.

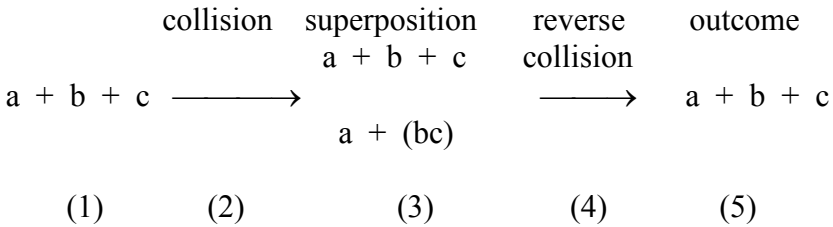
The condition for probabilistic collapse, formulated above, can readily be generalized to apply to more complicated and realistic inelastic interactions between “particles”.

According to the micro-realistic, fundamentally probabilistic version of quantum theory, indicated above, the state function, $\Phi(t)$, describes the actual physical state of the quantum system, from moment to moment. Quantum systems may be called “propensitons”. The physical (quantum) state of the propensiton evolves in accordance with Schrödinger’s time-dependent equation as long as the condition for a probabilistic transition to occur does not obtain. The moment it does obtain, the state jumps instantaneously and probabilistically, in the manner indicated above, into a new state. (All but one of a superposition of states, each with distinct “particles” associated with them, vanish.) The new state then continues to evolve in accordance Schrödinger’s equation until conditions for a new probabilistic transition arise.

Propensiton quantum theory (PQT), as we may call this micro-realistic, fundamentally probabilistic version of quantum theory, can recover all the experimental success of OQT. This follows from four points. First, OQT and PQT use the same dynamical equation, namely Schrödinger’s time-dependent equation. Secondly, whenever a position measurement is made, and a quantum system is detected, this invariably involves the creation of a new “particle” (bound or stationary system, such as the ionisation of an atom or the dissociation of a molecule, usually millions of these). This means that whenever a position measurement is made, the conditions for probabilistic transitions to occur, according to PQT, are satisfied. PQT will reproduce the predictions of OQT (given that PQT is provided with a specification of the quantum state of the measuring apparatus). Thirdly, all other observables of OQT, such as momentum, energy, angular momentum or spin, always involve (i) a preparation procedure which leads to distinct spatial locations being associated with distinct values of the observable to be measured, and (ii) a position measurement in one or other spatial location. This means that PQT can predict the outcome of measurements of all the

observables of OQT. Fourthly, insofar as the predictions of OQT and PQT differ, the difference is extraordinarily difficult to detect, and will not be detectable in any quantum measurement so far performed.

In principle, however, OQT and PQT yield predictions that differ for experiments that are extraordinarily difficult to perform, and which have not yet, to my knowledge, been performed. Consider the following evolution:-



Suppose the experimental arrangement is such that, if the superposition at stage (3) persists, then interference effects will be detected at stage (5). Suppose, now, that at stage (3) the condition for the superposition to collapse into one or other state, according to PQT, obtains. In these circumstances, OQT predicts interference at stage (5), whereas PQT predicts no interference at stage (5), (assuming the above evolution is repeated many times). PQT predicts that in each individual case, at stage (3), the superposition collapses probabilistically into one or other state. Hence there can be no interference.

3 Further Questions

It may be asked how $\Phi(t)$ can possibly represent the real physical state of a quantum system given that $\Phi(t)$ is a complex function of space and time. The answer is that $\Phi(t)$ can always be construed to depict two real functions of space and time.

It may be asked how $\Phi(t)$ can possibly represent the real physical state of a quantum system consisting of two (or more) quantum entangled “particles”, since in this case $\Phi(t)$ is a function of six dimensional configuration space plus time (or, in general, a function of $3N$ configuration space plus time, where N is the

number of quantum entangled “particles” that go to make up the system in question). In the case of two “particles”, we can construe $\Phi(r_1, r_2, t)$, where r_1 and r_2 are the spatial coordinates of “particles” 1 and 2 respectively, as depicting the propensity state of the system in real 3-dimensional physical space, as follows. $|\Phi(r_1, r_2, t)|^2 dV_1 dV_2$ represents the probability of the system interacting in a localizing (wave-packet-collapsing) way such that “particle” 1 interacts in volume element dV_1 about spatial coordinates r_1 , and “particle” 2 interacts in volume element dV_2 about spatial coordinates r_2 . The quantum entangled nature of the system means that as r_2 is changed, so the probability of “particle” 1 interacting in dV_1 about r_2 will, in general, change too.

It may be objected that postulate 1(A+B) provides no mechanism for quantum systems to be localized. This is not correct. If a highly localized system, S_1 , interacts inelastically with a highly unlocalized system, S_2 , in such a way that a probabilistic transition occurs, then S_1 will localize S_2 . If an atom or nucleus emits a photon which travels outwards in a spherical shell and which is subsequently absorbed by a localized third system, the localization of the photon will localize the emitting atom or nucleus with which it was quantum entangled.

Postulate 1(A+B) above has been formulated for rearrangement collisions. But the postulate is intended to apply to inelastic interactions that lead to the creation (or annihilation) of new particles, as in interactions such as $e^- + e^+ \rightarrow 2\gamma$. Such interactions require that one employs relativistic QT, which is beyond the scope of the present paper. It deserves to be noted, however, that the root idea that probabilistic transitions occur when new “particles” are created can be interpreted in a number of different ways.

(1) There is the option considered above. The inelastic interaction must be such that distinct “particle” channels have, associated with them, distinct asymptotic states which evolve in accordance with distinct Hamiltonians. This means at least that distinct “particles” have different masses associated with them (so that an excited state of a bound system is, potentially, a different

“particle” from the ground state, since the excited state will be slightly more massive than the ground state).

(2) As above, except that, for two interaction channels to differ it is not sufficient

that “particles” associated with the two channels have distinct masses; either there are different numbers of “particles” (counting a bound system as one “particle”) associated with different channels, or there is at least one “particle” which has a different charge, or force, associated with it.

(3) For a probabilistic transition to occur, rest mass must be converted into energy of “particles” without rest mass (eg photons), or vice versa.

(4) For a probabilistic transition to occur, fermions must be converted into bosons, or vice versa.

Only experiment can decide between these options. The import of this paper, and of previous papers published by the author (Maxwell, 1972b; 1973a; 1973b; 1976b; 1982; 1988; 1993c; 1994; 1995; 1998, ch. 7) is that a major research effort ought to get underway, both theoretical and experimental, devoted to exploring and testing rival collapse hypotheses. Only in this way will a version of quantum theory be developed free of the defects of OQT which also meets with greater empirical success than OQT. Only in this way will physics succeed in providing some kind of answer to the two childish elementary, inter-related questions with which we began.

4 Quantum Confusions a Part of a Historical Pattern

I conclude with a historical remark. I have argued that the long-standing failure to solve the mysteries of the quantum domain – and so to develop a fully acceptable version of quantum theory – is due to the misguided attempt to understand the probabilistic quantum domain in terms of deterministic metaphysical ideas appropriate to the earlier theories of classical physics. As a result of the failure to solve the wholly misguided traditional wave/particle problem, Heisenberg, Bohr, Born and others developed quantum theory as a theory about the results of performing *measurements*, which seemed successfully to avoid the

need to specify precisely the nature of quantum systems, but which unintentionally led to the creation of a theory with severe, if somewhat surreptitious, defects.

This pattern of confusion has occurred on at least two earlier occasions in the history of physics. On these occasions, too, physicists have attempted to interpret a new theory in terms of old, inappropriate metaphysics; the failure of this misguided effort then leads to despair at the possibility of interpreting the new theory realistically. It leads to instrumentalism, in other words, to the view that physical theories have to be interpreted as being about observable phenomena, and not about unobservable physical entities such as particles and fields. Eventually, however, the new theory may be interpreted in terms of new appropriate metaphysics. Physicists, one might say, are brilliant when it comes to equations, but not so brilliant – or at least very conservative – when it comes to metaphysics.

An example is Newton's theory of gravitation which postulates a force at a distance between bodies with mass. The reigning metaphysical idea at the time was the corpuscular hypothesis, the thesis that nature is made up of tiny corpuscles which interact only by contact. This thesis functioned as a standard of intelligibility: no fundamental physical theory could claim to be intelligible if it could not be interpreted in terms of the corpuscular hypothesis. The impossibility of interpreting Newton's theory of gravitation in terms of the corpuscular hypothesis initially led some of Newton's most eminent contemporaries to reject Newton's theory. Thus Huygens, in a letter to Leibniz, writes: "Concerning the Cause of the flux given by M. Newton, I am by no means satisfied [by it], nor by all the other Theories that he builds upon his Principle of Attraction, which seems to me absurd. . . I have often wondered how he could have given himself all the trouble of making such a number of investigations and difficult calculations that have no other foundation than this very principle" (Koyre, 1965, pp. 117-8). Newton in a sense agreed, as is indicated by his remark: "That gravity should be innate, inherent and essential to matter, so that one body may act upon another, at a distance through a vacuum, without the mediation of anything else. . . is to me so great an

absurdity, that I believe no man who has in philosophical matters a competent faculty of thinking can ever fall into it” (Burt, 1932, pp. 265-6). The impossibility of interpreting the law of gravitation in terms of the corpuscular hypothesis, in terms of action-by-contact, led Newton to interpret the law instrumentalistically, as specifying the way bodies move without providing any kind of explanation for the motion, in terms of unobservable forces. Subsequently, however, Boscovich and others were able to conceive of a metaphysical view more appropriate to Newton’s new theory, according to which nature is made up of point-particles, with mass, each point-particle being surrounded by a rigid, spherically-symmetric, centrally directed field of force which varies with distance. Reject the corpuscular hypothesis and adopt, instead, this new Boscovichean metaphysics, and Newton’s theory ceases to be incomprehensible, and becomes the very model of comprehensibility.

Another example is provided by James Clerk Maxwell’s theory of electrodynamics. Maxwell himself, and most of his contemporaries and immediate successors, sought to interpret the electromagnetic field in terms of a material substratum, the hypothetical aether, itself to be understood in Newtonian terms. A tremendous amount of effort was put into trying to understand Maxwell’s field equations in terms of the aether. Faraday, who appreciated that one should take the electromagnetic field as a new kind of physical entity, and explain matter in terms of the field rather than try to explain the field in terms of a kind of hypothetical matter (the aether), was ignored. The unrealistic character, and ultimate failure, of mechanical models of the electromagnetic field led many to hold that the real nature of the field must remain a mystery. The most that one could hope for from Maxwell’s equations, it seemed, was the successful prediction of observable phenomena associated with electromagnetism. This instrumentalistic attitude remained even after the advent of Einstein’s special theory of relativity in 1905, which might be interpreted as giving credence to the idea that it is the field that is fundamental. Gradually, however, Einstein and others came to adopt the view that one should see the field as a new kind of physical entity, quite distinct from corpuscle and point-particle.

There are two lessons to be learned from these episodes, one for quantum theory specifically, the other for theoretical physics in general. In the first place, quantum theory, if fundamentally probabilistic, needs to be formulated as a theory about fundamentally probabilistic physical entities – propensities – however weird these may seem given our common sense and classical intuitions. We require a fully micro-realistic version of quantum theory which, though testable, says nothing about “observables” or “measurement” in the basic postulates of the theory at all. Secondly, if theoretical physics is to free itself from the obstructive tendency to interpret new theories in terms of old, inappropriate metaphysics, physicists need to recognize that metaphysical ideas are inevitably an integral part of theoretical physics, and need to be developed and improved in the light of new theoretical developments. Elsewhere (Maxwell, 1998), I have argued that, in order to construe physics as a rational enterprise, we need to see physics as making a hierarchy of metaphysical assumptions concerning the comprehensibility and knowability of the universe, these assumptions becoming increasingly insubstantial, and thus increasingly likely to be true, as we ascend the hierarchy. According to this “aim-oriented empiricist” view, this hierarchy creates a framework of reasonably secure, permanent assumptions (and associated methods) within which much more specific and fallible assumptions (and associated methods), low down in the hierarchy, can be revised and improved. If ever the physics community came to accept and put into scientific practice this aim-oriented empiricist methodology, then the best available metaphysical ideas might lead the way to the discovery of new physical theories, instead of obstructing interpretation and understanding of theories that have been discovered (and thus also obstructing the discovery of new theories). In one exceptional case in the history of physics, the new metaphysics came first, led the way, and actually made possible the subsequent discovery of the new theory. This happened when Einstein discovered general relativity. Einstein first hit upon the metaphysical idea that gravitation is due to the curvature of space-time, and then subsequently discovered how to capture this idea precisely in the field equations of general relativity. In stark contrast to the cases

of Newtonian theory, Maxwellian classical electrodynamics and quantum theory, general relativity was discovered as a result of the prior development of new appropriate metaphysics, instead of the discovery of the new theory, if anything, being obstructed by current metaphysical ideas, the theory being misunderstood and misinterpreted by such ideas, once discovered. That Einstein's discovery of general relativity should stand out in this way is not, in my view, surprising: as I have argued elsewhere (Maxwell 1993a, pp, 275-305), Einstein both put into practice, and upheld, a conception of science close to that of aim-oriented empiricism.

References

- "A Blueprint for Survival", 1972, *The Ecologist*, vol. 2, No. 1.
- Abott, P. and Wallace, C., 1990, *An introduction to sociology: feminist perspectives*, Routledge, London.
- Aitchison, I. and A. Hey, 1982, *Gauge Theories in Particle Physics*. Bristol: Adam Hilger.
- Aspect, A., Grangier, P. and Roger, G., 1982, "Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell's Inequalities", *Physical Review Letters*, 48, pp. 91-4.
- Atkins, P. W., 1983, *Molecular Quantum Mechanics*, Oxford University Press, Oxford.
- Ayer, A. J., 1965, *The Problem of Knowledge*, Penguin Books, Middlesex.
- Ayer, A. J., 1969, *Metaphysics and Common Sense*, Macmillan, London.
- Ayer, A. J., *et al.*, 1967, *The Revolution in Philosophy*, Macmillan, London.
- Bacciagaluppi, G., 2003, "The Role of Decoherence in Quantum Mechanics", *The Stanford Encyclopedia of Philosophy (Winter 2003 Edition)*, Edward N. Zalta (ed.),
URL = <<http://plato.stanford.edu/archives/win2003/entries/qm-decoherence/>> .
- Barker, S. F., 1957, *Induction and Hypothesis: A Study of the Logic of Confirmation*, Cornell University Press, Ithaca.
- Barnes, B., 1974, *Scientific Knowledge and Sociological Theory*, Routledge & Kegan Paul, London.
- Barnes, B., D. Bloor and J. Henry, 1996, *Scientific Knowledge: A Sociological Analysis*, University of Chicago Press, Chicago.
- Bartelborth, T., 2002, "Explanatory Unification", *Synthese* 130, pp. 91-107.
- Bell, J. S., 1964, "On the Einstein Podolsky Rosen paradox", *Physics*, 1, pp. 195-200 (reprinted in Bell, 1987).
- Bell, J. S., 1987, *Speakable and unspeakable in quantum mechanics*, Cambridge University Press, Cambridge.

- Berlin, I., 1980, *Against the Current*, Hogarth Press, London.
- Berlin, I., 1999, *The Roots of Romanticism* Chatto and Windus, London.
- Bloor, D., 1976, *Knowledge and Social Imagery*, Routledge & Kegan Paul, London.
- Bohm, D., 1952, A Suggested Interpretation of the Quantum Theory in terms of "Hidden" Variables, *Physical Review* 85, 166-79 and 180-93.
- Bohm, D. and Y. Aharonov, 1957, "Discussion of Experimental Proof for the Paradox of Einstein, Rosen and Podolsky", *Physical Review*, 108, pp. 1070-6.
- Born, M., 1971, *The Born-Einstein Letters*, Macmillan.
- Boyd, R., 1980, 'Scientific Realism and Naturalistic Epistemology', *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, vol 2, University of Chicago Press, Chicago, pp. 613-662.
- Braithwaite, R., 1953, *Scientific Explanation*, Cambridge University Press, Cambridge.
- Brown, J. R., 2001, *Who Rules in Science?*, Harvard University Press, Cambridge, Mass.
- Buchdahl, G., 1969, *Metaphysics and the Philosophy of Science*, Blackwell, Oxford.
- Burt, E. A., 1932, *The Metaphysical Foundations of Modern Science*, Routledge and Kegan Paul, London (first published in 1924).
- Cartwright, N., 1999, *The Dappled World*, Cambridge University Press, Cambridge.
- Chakravartty, A., 1999, *Times Higher Educational Supplement*, 24 September, p. 24.
- Chart, D., 2000, *Philosophical Books* 41, no. 4, 2000, pp. 283-5.
- Chui, H. and Hoffmann, W.F., 1964, *Gravitation and Relativity*, Benjamin, New York.
- Collingridge, D., 1985, "Reforming Science", *Social Studies of Science* 15, Nov., pp. 763-69.
- Deutscher, M., 1968, "Popper's Problem of an Empirical Base", *Australasian Journal of Philosophy*, vol. 46, pp. 277-288.
- Duhem, P., 1962, *The Aim and Structure of Physical Theory*, Atheneum, New York.

- Dukas, H. and B. Hoffmann, eds., 1979, *Albert Einstein: The Human Side*, Princeton University Press, Princeton.
- Dupré, J., 1995, *The Disorder of Things*, Harvard University Press, Cambridge, Mass.
- Durant, J., 1997, "Beyond the Scope of Science", *Science and Public Affairs*, Spring, pp. 56-7.
- Easlea, B., 1973, *Liberation and the Aims of Science: An Essay on Obstacles to the Building of a Beautiful World*, Chatto & Windus, London.
- Easlea, B., 1986, "Review of From Knowledge to Wisdom", *Journal of Applied Philosophy* 3, pp. 139-40.
- Einstein, A., 1905, "Zur Elektrodynamik bewegter Körper", *Annalen der Physik*, 17, pp. 891-921.
- Einstein, A., 1969, "Autobiographical Notes" and "Replies to Criticisms", in Schilpp (1969), pp. 1-94
- Einstein, A., 1973, *Ideas and Opinions*, Souvenir Press, London.
- Einstein, A., Podolsky, B. and Rosen, N., 1935, "Can Quantum-mechanical Description of Reality be Considered Complete?", *Physical Review* 47, pp. 777-80.
- Einstein, A., Lorentz, H. A., Minkowski, H. and Weyl, H., 1952, *The Principle of Relativity*, Dover.
- Emmerson, J., 1972, *Symmetry Principles in Particle Physics*, Clarendon Press, Oxford.
- Fargaus, J., ed., 1993, *Readings in Social Theory*, McGraw-Hill, New York.
- Feyerabend, P., 1965, "Problems of Empiricism", in *Beyond the Edge of Certainty*, edited by R. Colodny, *University of Pittsburgh Series in the Philosophy of Science*, vol. 2, Prentice-Hall, Englewood Cliffs, pp. 145-260.
- Feyerabend, P., 1968-69, "On a Recent Critique of Complementarity," *Philosophy of Science*, vol. 35, pp. 309-31; and vol. 36, pp. 82-105.
- Feyerabend, P., 1970, "Problems of Empiricism II", in *The Nature and Function of Scientific Theories*, edited by R. Colodny, *University of Pittsburgh series in the Philosophy of Science*, vol. 4, University of Pittsburgh Press, Pittsburgh.
- Feyerabend, P., 1978, *Against Method*, Verso, London.
- Feyerabend, P., 1987, *Farewell to Reason*, Verso, London.

- Feynman, R., R. Leighton and M. Sands, 1965, *The Feynman Lectures on Physics*, vol. II, Addison-Wesley, Reading, Mass.
- Fine, A., 1986, *The Shaky Game*, University of Chicago Press, Chicago.
- Franklin, A., 1986, *The Neglect of Experiment*, Cambridge University Press, Cambridge.
- Franklin, A., 1990, *Experiment: Right or Wrong*, Cambridge University Press, Cambridge.
- French, A. P., ed., 1979, *Einstein: A Centenary Volume*, Heinemann.
- Friedman, M., 1974, "Explanation and Scientific Understanding", *Journal of Philosophy* 71, pp. 5-19.
- Friedman, M., 1983, *Foundations of Space-Time Theories*, Princeton University Press.
- Fromm, E., 1960, *The Fear of Freedom*, Routledge & Kegan Paul, London.
- Fromm, E., 1963, *The Sane Society*, Routledge & Kegan Paul, London.
- Gascardi, A., 1999, *Consequences of Enlightenment*, Cambridge University Press, Cambridge.
- Gay, P., 1973, *The Enlightenment: An Interpretation*, Wildwood House, London.
- Ghirardi, G.C., Rimini, A. and Weber, T., 1986, Unified Dynamics for Microscopic and Macroscopic Systems, *Physical Review D* 34, 470-91.
- Ghirardi, G. C., and A. Rimini, 1990, Old and New Ideas in the Theory of Quantum Measurement, in *Sixty-Two Years of Uncertainty*, edited by A. Miller, Plenum Press, New York, 167-91.
- Gillespie, T. D., 1973, *A Quantum Mechanical Primer*, International Textbook, Aylesbury.
- Glymour, C., 1980, *Theory and Evidence*, Princeton University Press, Princeton.
- Goodman, N., 1954, *Fact, Fiction and Forecast*. London: Athlone Press.
- Goodman, N., 1972, *Problems and Projects*. New York: Bobbs-Merrill.

- Gray, J. G., 1972, *The Promise of Wisdom*, Harper & Row, New York.
- Griffiths, D., 1987, *Introduction to Elementary Particles*, Wiley, New York:
- Gross, P. and N. Levitt, 1994, *Higher Superstition: The Academic Left and Its Quarrels with Science*, John Hopkins University Press, Baltimore.
- Gross, P., N. Levitt and M. Lewis, eds., 1996, *The Flight from Science and Reason*, John Hopkins University Press, Baltimore.
- Grünbaum, A., 1963, *Philosophical Problems of Space and Time*, Routledge and Kegan Paul, London.
- Haberman, F. W., 1972, *Nobel Lectures, Peace, 1951-1970*, Vol. 3, Elsevier Publishing Co., Amsterdam.
- W. O. Hagstrom, W. O., 1965, *The Scientific Community*, Basic Books, New York.
- Harding, S., 1986, *The Feminist Question in Science*, Open University Press, Milton Keynes.
- Harré, R., 1986, *Varieties of Realism*, Blackwell, Oxford.
- Harris, M., 1978, *Cannibals and Kings*, Fontana, London.
- Hayek, F., 1979, *The Counter-Revolution of Science*, Liberty Press, Indianapolis.
- Hempel, C. G., 1965, *Aspects of Scientific Explanation*. New York: Free Press.
- Higgins, R., 1978, *The Seventh Enemy: The Human Factor in the Global Crisis*, Hodder & Stoughton, London.
- Hodgson, D., 2002, "Review of The Human World in the Physical Universe", *Journal of Consciousness Studies* 9, pp. 93-4.
- Hoffmann, B., 1973, *Albert Einstein: Creator and Rebel*, Hart-Davis and MacGibbon, London.
- Holton, G., 1973, *Thematic Origins of Modern Science*, Harvard University Press, Cambridge, Mass.
- Hooker, G. A., 1987, *A Realistic Theory of Science*, State University of New York Press, Albany.
- Howson, C., 2000, *Hume's Problem*, Oxford University Press, Oxford.
- Howson, C. and Urbach, P., 1993, *Scientific Reasoning*, Open Court, La Salle.

- Hull, D. L., 1988, *Science as a Process*, University of Chicago Press, Chacago.
- Hume, D., 1959, *A Treatise of Human Nature*. London: J. M. Dent and Sons.
- Huxley, A., 1938, *Means and Ends: An Enquiry into the Nature of Ideals and into the Methods employed for their Realization*, Chatto & Windus, London.
- Isham, C., 1989, *Lectures on Groups and Vector-Spaces for Physicists*, World Scientific, London.
- Jeffreys, H., 1957, *Scientific Inference*. Cambridge: Cambridge University Press.
- Jeffries, H. and D. Wrinch, 1921, "On Certain Fundamental Principles of Scientific Enquiry", *Philosophical Magazine* 42, pp. 269-98.
- Jones, H., 1990, *Groups, Representations and Physics*, Adam Hilger, Bristol.
- Juhl, C. F., 2000, *International Philosophical Quarterly*, XL, No. 4, pp. 517-8.
- Kant, I., 1961, *Critique of Pure Reason*. Translated by N. K. Smith. London: Macmillan and Co.
- Kitcher, P., 1981, "Explanatory Unification", *Philosophy of Science* 48, pp. 507-31.
- Kitcher, P., 1993, *The Advancement of Science*, Oxford University Press, New York.
- Kneller, G. F., 1978, *Science as a Human Endeavor*, Columbia University Press, New York. pp. 84 and 91.
- Koertge, N., ed., 1998, *A House Built on Sand*, Oxford University Press, Oxford.
- A. Koyre, 1965, *Newtonian Studies*, Chapman and Hall.
- Koyré, A., 1973, *Metaphysics and Measurement*. London: Methuen.
- Kneller, G, F., 1978, *Science as a Human Endeavor*, Columbia University Press, New York.
- S. Kripke, 1981, *Naming and Necessity*, Blackwell, Oxford.
- Kuhn, T. S., 1970a, *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press (first published 1962).

- Kuhn, T. S., 1970b, "Logic of Discovery or Psychology of Research?" and "Reflections on my Critics", in Lakatos and Musgrave (1970), pp. 1-23 and 231-278.
- Kuhn, T. S., 1977, *The Essential Tension*, University of Chicago Press, Chicago.
- Lakatos, I., 1968a, "The Changing Problem of Inductive Logic" in L. Lakatos, ed., *The Problem of Inductive Logic*, North Holland, Amsterdam, pp. 315-417.
- Lakatos, I., 1968b, "Criticism and the Methodology of Scientific Research Programmes." *Proceedings of the Aristotelian Society*, 69, 149-186.
- Lakatos, I., 1970, "Falsification and the Methodology of Scientific Research Programmes", in Lakatos and Musgrave (1970), pp. 91-195.
- Lakatos, I., 1971, "History and its Rational Reconstructions." In *Boston Studies in the Philosophy of Science*. Vol. 8. Edited by R. C. Buck and R. S. Cohen. Dordrecht: D. Reidel Publishing Co.
- Lakatos, I., 1978, *The Methodology of Scientific Research Programmes*, ed. J. Worrall and G. Currie, Cambridge University Press, Cambridge.
- Lakatos, I., and Musgrave, A., eds., 1970, *Criticism and the Growth of Knowledge*, Cambridge University Press, Cambridge.
- Lanczos, C., 1974, *The Einstein Decade (1905-1915)*, Paul Elek (Scientific Books) Ltd.
- Latour, B., 1987, *Science in Action*, Open University Press, Milton Keynes.
- Laudan, L., 1977, *Progress and Its Problems*, University of California Press, Berkeley.
- Laudan, L., 1980, "A Confutation of Convergent Realism", *Philosophy of Science* 48, pp. 19-48.
- Laudan, L., 1984, *Science and Values*, University of California Press, Berkeley.
- Laudan, L., 1987, "Progress or Rationality? The Prospects for Normative Naturalism", *American Philosophical Quarterly* 24, pp. 19-31.
- Lanski, G. et al., 1995, *Human Societies: An Introduction to Macrosociology*, McGraw-Hill, New York.

- Lipton, P., 2004, *Inference to the Best Explanation*, 2nd edition, Routledge, London.
- Longuet-Higgins, L., 1984, "For goodness sake", *Nature* 312, 15 Nov., p.204.
- Mach, E., "The Guiding Principles of My Scientific Theory of Knowledge and Its Reception by My Contemporaries." In *Physical Reality*. Edited by S. Toulmin. New York: Harper Torchbooks. 1970.
- MacIntyre, A., 2009, *British Journal of Educational Studies* 57, Issue 4, p. 358.
- Mandl, F. and G. Shaw, 1984, *Quantum Field Theory*, Wiley, New York.
- Mannheim, K., 1952, *Essays in the Sociology of Knowledge*, Routledge & Kegan Paul, London.
- Mathias, P., ed., 1972, *Science and Society 1600-1900*, Cambridge University Press, London.
- Macionis, J. and Plummer, K., 1997, *Sociology: A Global Introduction*, Prentice Hall, New York.
- Maudlin, T., 1996, "On the Unification of Physics", *Journal Of Philosophy* 93, pp. 129-44.
- Maxwell, N., 1968, "Can there be Necessary Connections between Successive Events?" *British Journal for the Philosophy of Science* 19, 1-25.
- Maxwell, N., 1972a, "A Critique of Popper's Views on Scientific Method." *Philosophy of Science* 39, 131-152.
- Maxwell, N., 1972b, "A New Look at the Quantum Mechanical Problem of Measurement", *American Journal of Physics*, 40, pp. 1431-5.
- Maxwell, N., 1973a, "Alpha Particle Emission and the Orthodox Interpretation of Quantum Mechanics", *Physics Letters*, 43A, pp. 29-30.
- Maxwell, N., 1973b, "The Problem of Measurement – Real or Imaginary?", *American Journal of Physics*, 41, pp. 1022-5.
- Maxwell, N., 1974, "The Rationality of Scientific Discovery, Parts 1 and II", *Philosophy of Science* 41, pp. 123-53 and 247-95.
- Maxwell, N., 1975, "Does the Minimal Statistical Interpretation of Quantum Mechanics Resolve the Measurement Problem?", *Methodology and Science* 8, pp. 84-101.

- Maxwell, N., 1976a, *What's Wrong With Science? Towards a People's Rational Science of Delight and Compassion*, Bran's Head Books, Hayes (2nd edition, Pentire Press, London).
- Maxwell, N., 1976b, "Towards a Micro Realistic Version of Quantum Mechanics", *Foundations of Physics* 6, 1976, pp. 275-92 and 661-76.
- Maxwell, N., 1977a, "Articulating the Aims of Science", *Nature*, vol. 265, p. 2.
- Maxwell, N., 1977b, 'Science and Values', *Times Higher Educational Supplement*, 4 November, p. 27.
- Maxwell, N., 1979, "Induction, Simplicity and Scientific Progress", *Scientia*, vol. 114, pp. 629-53 (Italian translation, pp. 655-74.)
- Maxwell, N., 1980, "Science, Reason, Knowledge and Wisdom: A Critique of Specialism", *Inquiry* 23, pp. 19-81.
- Maxwell, N., 1982, "Instead of Particles and Fields: A Micro Realistic Quantum "Smearon" Theory", *Foundations of Physics* 12, pp. 607-31.
- Maxwell, N., 1984a, *From Knowledge to Wisdom: A Revolution in the Aims and Methods of Science*, Blackwell, Oxford.
- Maxwell, N., 1984b, "Guiding Choices in Scientific Research", *Bulletin of Science, Technology and Society* 4, pp. 316-34.
- Maxwell, N., 1985a, "Are Probabilism and Special Relativity Incompatible?", *Philosophy of Science* 52, pp. 23-43.
- Maxwell, N., 1985b, "From Knowledge to Wisdom: the Need for an Intellectual Revolution", *Science, Technology and Society Newsletter* 21, pp. 55-63.
- Maxwell, N., 1986, "The Fate of the Enlightenment: Reply to Kekes", *Inquiry* 29, pp. 79-82.
- Maxwell, N., 1987, "Wanted: a new way of thinking", *New Scientist* 14, May, p. 63.
- Maxwell, N., 1988, "Quantum Propensiton Theory: A Testable Resolution of the Wave/Particle Dilemma", *The British Journal for the Philosophy of Science* 39, pp. 1-50.
- Maxwell, N., 1991, "How Can We Build a Better World?" in J. Mittelstrass, ed., *Einheit der Wissenschaften*, Walter de Gruyter, pp. 388-427.

- Maxwell, N., 1992a, "What Kind of Inquiry Can Best Help Us Create a Good World?", *Science, Technology and Human Values* 17, pp. 205-27.
- Maxwell, N., 1992b, "What the Task of Creating Civilization has to Learn from the Success of Modern Science: Towards a New Enlightenment", *Reflections on Higher Education* 4, pp. 47-69.
- Maxwell, N., 1993a, "Induction and Scientific Realism: Einstein versus van Fraassen", *British Journal for the Philosophy of Science*, vol. 44, pp. 61-79, 81-101, and 275-305.
- Maxwell, N., 1993b, "Does Quantum Mechanics Undermine, or Support, Scientific Realism?", *The Philosophical Quarterly* 43, pp. 139-57.
- _____, 1993c, Beyond Fapp: Three Approaches to Improving Orthodox Quantum Theory and An Experimental Test, in *Bell's Theorem and the Foundations of Modern Physics*, edited by A. van der Merwe, F. Selleri and G. Tarozzi, World Scientific, 362-70.
- Maxwell, N., 1994, "Particle Creation as the Quantum Condition for Probabilistic Events to Occur", *Physics Letters A* 187, pp. 351-355.
- Maxwell, N., 1994b, "Towards a New Enlightenment: What the task of Creating Civilization has to learn from the Success of Modern Science", in *Academic Community: Discourse or Discord?*, R. Barnett (ed.), Jessica Kingsley, pp. 86-105.
- Maxwell, N., 1995, "A Philosopher Struggles to Understand Quantum Theory: Particle Creation and Wavepacket Reduction", in *Fundamental Problems in Quantum Physics*, ed. M. Ferrero and A. van der Merwe, Kluwer Academic, London, pp. 205-14.
- Maxwell, N., 1997a, "Must Science Make Cosmological Assumptions if it is to be Rational?", in *The Philosophy of Science: Proceedings of the Irish Philosophical Society Spring Conference*, ed., T. Kelly, Irish Philosophical Society, Maynooth, pp. 98-146.
- Maxwell, N., 1997b, "Science and the environment: A new enlightenment", *Science and Public Affairs*, Spring, pp. 50-56.
- Maxwell, N., 1998, *The Comprehensibility of the Universe: A New Conception of Science*, Oxford University Press, Oxford.

- Maxwell, N., 1999a, "Has Science Established that the Universe is Comprehensible?", *Cogito*, vol. 13, pp. 139-145.
- Maxwell, N., 1999b, "Are There Objective Values?", *The Dalhousie Review* 79, pp. 301-317.
- Maxwell, N., 2000a, "Can Humanity Learn to become Civilized? The Crisis of Science without Civilization", *Journal of Applied Philosophy* 17, pp. 29-44.
- Maxwell, N., 2000b, "A new conception of science", *Physics World* 13, No. 8, pp. 17-18.
- Maxwell, N., 2001a, *The Human World in the Physical Universe*, Rowman and Littlefield, Lanham, Maryland.
- Maxwell, N., 2001b, "Can Humanity Learn to Create a Better World? The Crisis of Science without Wisdom", in *The Moral Universe*, ed. T. Bentley and D. Stedman Jones, *Demos Collection* 16, pp. 149-156.
- Maxwell, N., 2001c, "Weinert's Review of *The Comprehensibility of the Universe*", *Philosophy* 76, 2001, pp. 297-303
- Maxwell, N., 2002a, "Is Science Neurotic?", *Metaphilosophy*, vol. 33, pp. 259-99.
- Maxwell, N., 2002b, "The Need for a Revolution in the Philosophy of Science", *Journal for General Philosophy of Science*, vol. 33, pp. 381-408.
- Maxwell, N., 2002c, "Cutting God in Half", *Philosophy Now* 35, March/April, pp. 22-25.
- Maxwell, N., 2002d, "Is Science Neurotic? *Metaphilosophy* 33, pp. 259-299.
- Maxwell, N., 2002e, "Karl Raimund Popper", in *British Philosophers, 1800-2000*, P. Dematteis, P. Fosl and L. McHenry (eds.), Brucoli Clark Layman, Columbia.
- Maxwell, N., 2003a, Two Great Problems of Learning, *Teaching in Higher Education* 8, pp. 129-134.
- Maxwell, N., 2003b, Science, Knowledge, Wisdom and the Public Good, *Scientists for Global Responsibility Newsletter* 26, pp. 7-9.
- Maxwell, N., 2003c, Do Philosophers' Love Wisdom, *The Philsophers' Magazine*, Issue 22, 2nd quarter, pp. 22-24.

- Maxwell, N., 2003d, "The Need for a Revolution in the Philosophy of Science", *Journal for General Philosophy of Science* 34, pp. 1-28.
- Maxwell, N., 2004, *Is Science Neurotic?*, Imperial College Press, London.
- Maxwell, N., 2006, "Aim-Oriented Empiricism: David Miller's Critique", *PhilSci Archive*, <http://philsci-archive.pitt.edu/3092/>.
- Maxwell, N., 2007, *From Knowledge to Wisdom: A Revolution for Science and the Humanities*, Pentire Press, London (2nd edition, revised and extended, of Maxwell, 1984).
- Maxwell, N., 2008, "Do We Need a Scientific Revolution?", *Journal for Biological Physics and Chemistry*, vol. 8, no. 3, September pp. 95-105.
- Maxwell, N., 2009a, "Are Universities Undergoing an Intellectual Revolution?", *Oxford Magazine*, No. 290, Eighth Week, Trinity Term, June 2009, pp. 13-16. (www.nick-maxwell.demon.co.uk/Essays.htm#essayiiia)
- Maxwell, N., 2009b, "Muller's Critique of the Argument for Aim-Oriented Empiricism", *Journal for General Philosophy of Science* 40, pp. 103-114.
- Maxwell, N., 2009c, "Replies and Reflections" in McHenry (2009b, pp. 249-313).
- Maxwell, N., 2010a, *Cutting God in Half – And Putting the Pieces Together Again: A New Approach to Philosophy*, Pentire Press, London.
- Maxwell, N., 2010b, "Reply to Comments on *Science and the Pursuit of Wisdom*", *Philosophia* 38, pp. 667-90
- Maxwell, N., 2011a, "Is the Quantum World Composed of Propensitons?", in *Probabilities, Causes and Propensities in Physics*, edited by Mauricio Suárez, Synthese Library, Springer, Dordrecht, pp. 219-241.
- Maxwell, N., 2011b, "A Priori Conjectural Knowledge in Physics", in *What Place for the A Priori?*, edited by Michael Shaffer and Michael Veber, Open Court, Chicago, pp. 211-240.
- Maxwell, N., 2013, "Has Science Established that the Cosmos is Physically Comprehensible?", to be published in *Recent Advances in Cosmology*, Nova Publishers Inc, New York: <http://discovery.ucl.ac.uk/view/people/ANMAX22.date.html>.

- Maxwell, N., 2012, "Popper's Paradoxical Pursuit of Natural Philosophy", in *Cambridge Companion to Popper*, edited by J. Shearmur and G. Stokes, Cambridge University Press, Cambridge.
- Maxwell, N., *The Aims of Science* (unpublished).
- McAllister, J., 1989, Truth and Beauty in Scientific Reason, *Synthese* 78, pp. 25-51.
- McAllister, J., 1990, Dirac and the Aesthetic Evaluation of Theories, *Methodology and Science* 23, pp. 87-102.
- McAllister, J., 1991, The Simplicity of Theories: Its Degree and Form, *Journal for General Philosophy of Science* 22, pp. 1-14.
- McAllister, J. W., 1996, *Beauty and Revolution in Science*, Cornell University Press, Ithaca.
- McHenry, L., 2000, *Mind* 109, 2000, pp. 162-166.
- McHenry, L., 2009a, "Popper and Maxwell on Scientific Progress" in McHenry (2009b), pp. 233-47).
- McHenry, L., ed., 2009b, *Science and the Pursuit of Wisdom: Studies in the Philosophy of Nicholas Maxwell*, Ontos Verlag, Frankfurt.
- Mead, M., 1943, *Coming of Age in Samoa: a Study of Adolescence and Sex in Primitive Societies*, Penguin Books Ltd., Harmondsworth.
- Mellor, D.H., 1991, *Matters of Metaphysics*, Cambridge University Press, Cambridge, pp. 254-68.
- Merton, R. K., 1970, *Science, Technology and Society in Seventeenth-Century England*, Harper & Row, New York.
- Meyerson, E., 1930, *Identity and Reality*. London: George Allen and Unwin.
- Midgley, M., 1979, *Beast and Man: The Roots of Human Nature*, Harvester Press, Sussex.
- Midgley, M., 1986, "Is Wisdom Forgotten?", *University Quarterly: Culture, Education and Society* 40.
- Miller, D., 1974, "Popper's Qualitative Theory of Verisimilitude", *British Journal for the Philosophy of Science* 25, pp. 166-177.
- Miller, D., 2006, *Out of Error*, Ashgate, London.
- Misner, C. W., Thorne, K. and Wheeler, J. A., 1973, *Gravitation*, W. H. Freeman.
- Moriyasu, K., 1983, *An Elementary Primer for Gauge Theory*, World Scientific, Singapore.

- Morrison, M., 2000, *Unifying Scientific Theories*, Cambridge University Press, Cambridge.
- Mulkay, M., 1979, *Science and the Sociology of Knowledge*, George Allen & Unwin, London.
- Muller, F. A., 2004, Maxwell's Lonely War, *Studies in History and Philosophy of Modern Physics* 35, 2004, pp. 109-110 and p. 117.
- Muller, F. A., 2008, "In defence of constructivism: Maxwell's master argument and aberrant theories, *Journal of General Philosophy of Science*, vol. 39, pp. 131-56.
- Musgrave, A., 1993, *Common Sense, Science and Scepticism*, Cambridge University Press, Cambridge.
- Newton, I., 1962, *Principia*, vol. II, trans. A. Motte and F. Cajori, California University Press, California (first published 1687).
- Newton-Smith, W., 1981, *The Rationality of Science*, Routledge and Kegan Paul, London.
- Nola, R. and H. Sankey, 2000, "A Selective Survey of Theories of Scientific Method", in R. Nola and H. Sankey (eds.) *After Popper, Kuhn and Feyerabend*, Kluwer, Dordrecht.
- North, J., 1965, *The Measure of the Universe*, Clarendon Press, Oxford.
- Norton, J., 1984, "How Einstein Found His Field Equations: 1912-1915", *Historical Studies in the Physical Sciences* 14, pp. 253-316.
- Pais, A., 1982, *Subtle is the Lord. . .*, Oxford University Press, Oxford.
- Pandit, G. L., 2010, "How Simple is it for Science to Acquire Wisdom According to its Chicest Aims?", *Philosophia*, vol. 38, pp. 649-66.
- Penrose, R. and M. MacCallum, 1973, "Twistor Theory: An Approach to the Quantization of Fields and Space-Time", *Physics Reports*, 6C (No. 4).
- Penrose, R., 1986, Gravity and State Vector Reduction, in Isham, C. and Penrose, R., eds., *Quantum Concepts in Space and Time*, Oxford University Press, Oxford, 129-46.
- Pickering, A., 1984, *Constructing Quarks*, University of Chicago Press, Chicago.
- Poincaré, H., 1952, *Science and Hypothesis*, Dover, New York.

- Poincaré, H., Quoted in J. Hildebrand, *Science in the Making*. New York: Columbia University Press, 1957.
- Popper, K. R., 1957, "The Aim of Science", *Ratio*, vol. 1, pp. 24-35.
- Popper, K. R., 1957b, The Propensity Interpretation of the Calculus of Probability and the Quantum Theory, in *Observation and Interpretation*, edited by S. Körner, Butterworth, London, 65-70.
- Popper, K. R., 1959, *The Logic of Scientific Discovery*. London: Hutchinson.
- Popper, K. R., 1961, *The Poverty of Historicism*, Routledge and Kegan Paul, London (first published 1957).
- Popper, K. R., 1963, *Conjectures and Refutations*. London: Routledge and Kegan Paul.
- Popper, K. R., 1966a, *The Open Society and Its Enemies*, Routledge and Kegan Paul, London (first published 1945).
- Popper, K. R., 1966b, "A Theorem on Truth Content", in P. Feyerabend and G. Maxwell, eds., *Mind, Matter, and Method*. Minneapolis: University of Minnesota Press, Minneapolis.
- Popper, K. R., 1967, Quantum Mechanics without "The Observer", in *Quantum Theory and Reality*, edited by M. Bunge, Springer, Berlin, 7-44.
- Popper, K. R., 1970, "Normal Science and its Dangers", in Lakatos and Musgrave (1970), pp. 51-8.
- Popper, K. R., 1972, *Objective Knowledge*, Oxford University Press, Oxford.
- Popper, K. R., 1974, "Replies to Critics", in Schilpp (1974, vol. 2, pp. 961-1197).
- Popper, K. R., 1976a, *Unended Quest*, Fontana/Collins, London.
- Popper, K. R., 1976b, "Reason or Revolution?", in *The Positivist Dispute in German Sociology*, Heinemann Educational Books Ltd., London.
- Popper, K. R., 1977, 'How I see Philosophy', in A. Mercier *et al.* eds., *Philosophers on their Own Work*, vol. 3, Peter Lang, Las Vegas.
- Popper, K. R., 1982, *The Open Universe*, vol. 2 of *The Postscript to The Logic of Scientific Discovery*, Hutchinson, London.
- Popper, K. R., 1982a, *Quantum Physics and the Schism in Physics*, vol. 3 of *The Postscript*, Hutchinson, London.

- Popper, K. R., 1983, *Realism and the Aim of Science*, vol. 1 of *The Postscript*, Hutchinson, London.
- Popper, K. R. and J. Eccles, 1977, *The Self and Its Brain*, Springer-Verlag, London.
- Price, D. de S., 1961, *Science since Babylon*, Yale University Press, New Haven.
- Przibram, K., ed., 1967, *Letters on Wave Mechanics*, Vision Press Ltd., London.
- Rae, A. I. M., 2002, *Quantum Mechanics*, Institute of Physics Publishing, Bristol.
- Reichenbach, H., 1938, *Experience and Prediction*, Chicago University Press, Chicago.
- Reichenbach, H., 1958, *The Rise of Scientific Philosophy*. Berkeley: University of California Press.
- Riemann, G., 1929, "On the Hypotheses which Lie at the Foundation of Geometry", in D.E. Smith, ed., *A Source Book in Mathematics*, McGraw-Hill, London.
- Rescher, N., 1973, *The Primacy of Practice*, Blackwell, Oxford.
- Rescher, N., 1977, *Methodological Pragmatism*, Blackwell, Oxford.
- Rescher, N., 1987, *Scientific Realism*, Reidel, Dordrecht.
- Rogers, K., 2009b, "Metaphysics and Methodology: Aim-Oriented Empiricism", in McHenry (2009, pp. 217-232).
- Rose, H. and S. Rose, eds., 1976, *The Radicalization of Science*, Macmillan, London.
- Roush, S., 2001, *The Philosophical Review* 110, pp. 85-7.
- Russell, B., 1948, *Human Knowledge: Its Scope and Limits*, Allen and Unwin, London.
- Rudner, R. S., 1961, "An Introduction to Simplicity." *Philosophy of Science* 28, 109-119.
- Ryle, G., 1949, *The Concept of Mind*, Hutchinson, London.
- Ryle, G., 1967, 'Introduction', in Ayer (1967, pp. 1-11).
- Salaman, E., 1979, 'Memories of Einstein', *Encounter*, April issue.
- Salmon, W., 1968, "The Justification of Inductive Rules of Inference", in I. Lakatos, ed., *The Problem of Inductive Logic*, North Holland.
- Salmon, W., 1989, *Four Decades of Scientific Explanation*, University of Minnesota Press, Minneapolis.

- Scheffler, I., 1963, *The Anatomy of Inquiry*. New York: Alfred A. Knopf.
- Schilpp, P., ed., 1969, *Albert Einstein: Philosopher-Scientist*, Open Court, LaSalle, Illinois.
- Schilpp, P. A., ed., 1974, *The Philosophy of Karl Popper*, Open Court, La Salle, Illinois.
- Schumacher, E. F., 1973, *Small is Beautiful: A Study of Economics as if People Mattered*, Blond & Briggs, London.
- Segerstrale, U., ed., 2000, *Beyond the Science Wars*, State University of New York Press, Albany.
- Schiff, M., 2005, *Choice*, vol. 42, no. 2.
- Schurz, G., 1999, "Explanation as Unification", *Synthese* 120, pp. 95-114.
- Schutz, B. F., 1988, *A First Course in General Relativity*, Cambridge University Press, Cambridge.
- Shankland, R., 1963, "Conversations with Albert Einstein", *American Journal of Physics*, 31, pp. 37-47.
- Shanks, N., 2000, *Metascience* 9, 2000, pp. 294-8.
- Shapin, S., 1994, *A Social History of Truth*, University of Chicago Press, Chicago.
- Singer, I. B., 1975a, *The Manor*, Penguin Books, Harmondsworth.
- Singer, I. B., 1975b, *The Estate*, Penguin Books, Harmondsworth.
- Singer, I. B., 1975c, *A Friend of Kafka*, Penguin Books, Harmondsworth.
- Singer, I. B., 1977a, *The Séance and Other Stories*, Penguin Books, Harmondsworth.
- Singer, I. B., 1977b, *A Crown of Feathers and Other Stories*, Penguin Books, Harmondsworth.
- Singer, I. B., 1977c, *Enemies: A Love Story*, Penguin Books, Harmondsworth.
- Smart, J. J. C., 1963, *Philosophy and Scientific Realism*, Routledge and Kegan Paul, London.
- Smart, J. J. C., 2000, The Comprehensibility of the Universe, *British Journal for the Philosophy of Science* 51, 2000, pp. 907-911.
- Sokal, A., 2008, *Beyond the Hoax*, Oxford University Press, Oxford.

- Sokal, A. and J. Bricmont, 1998, *Intellectual Impostures*, Profile Books, London.
- Squires, E., 1986, *The Mystery of the Quantum World*, Adam Hilger, Bristol.
- Stachel, J., 1980, "Einstein and the Rigidly Rotating Disk", in A. Held, ed., *General Relativity and Gravitation 1*, Plenum Press, pp. 3-15.
- Stachel, J., ed., 1998, *Einstein's Miraculous Year: Five Papers That Changed the Face of Physics*, Princeton University Press, Princeton.
- Swinburne, R. G., 1970, "Choosing Between Confirmation Theories." *Philosophy of Science* 28, 602-613.
- Swinburne, R. G. (ed.), 1974, *The Justification of Induction*. London: Oxford University Press.
- Teich, M. and R. M. Young, eds., 1873, *Changing Perspectives in the History of Science*, Heinemann, London.
- Tichy, P., 1974, "On Popper's Definition of Verisimilitude", *British Journal for the Philosophy of Science* 25, pp. 155-160.
- Tischler, H., 1996, *Introduction to Sociology*, Harcourt Brace, Orlando.
- Turnbull, C., 1976, *The Forest People*, Picador, London.
- van Fraassen, B., 1980, *The Scientific Image*, Clarendon Press, Oxford.
- van Fraassen, B., 1985, "Empiricism in the Philosophy of Science", in P. M. Churchland & C. A. Hooker, eds., *Images of Science*, University of Chicago Press, Chicago.
- Vicente, A., 2010, "An Enlightened Revolt: on the Philosophy of Nicholas Maxwell", *Philosophia*, vol. 38, pp. 631-48.
- Worrall, J., 1989, "Why Both Popper and Watkins Fail to Solve the Problem of Induction" in *Freedom and Rationality: Essays in Honour of John Watkins*, Kluwer, Dordrecht.
- Watkins, J., 1958, "Confirmable and Influential Metaphysics", *Mind* 67, pp. 344-365.
- Watkins, J., 1984, *Science and Scepticism*, Princeton University Press, Princeton.
- Weber, E., 1999, "Unification", *Synthese* 118, pp. 479-99.
- Weinert, F., 2000, *Philosophy* 75, pp. 296-309.
- Werskey, G., 1978, *The Visible College*, Allen Lane, London.

- Weyl, H., 1963, *Philosophy of Mathematics and Natural Science*.
New York: Athenium (first published in German in 1926).
- Wheeler, J. A., 1968 “Superspace and the Nature of Quantum
Geometrodynamics”, in C. M. DeWitt and J. A. Wheeler, eds.,
Battelle Recontre, Benjamin, London.
- Wheeler, J. A., 1972, *Geometrodynamics*, Academic Press,
London.
- Whiteley, C. H., 1955, *An Introduction to Metaphysics*, Methuen,
London.
- Whitrow, G., 1973, *Einstein: The Man and His Achievement*,
Dover Press,
- Wigner, E.P., 1970, *Symmetries and Reflections*, M.I.T. Press,
London.
- Wu, C. S. *et al.*, 1957, *Physical Review* 105, p. 1413.
- Zahar. E., 1973, “Why did Einstein's Programme Supersede
Lorentz's?”, *British Journal for the Philosophy of Science*, 24,
pp. 95-123.
- Zamyatin, Y., 1972, *We*, Penguin, Harmondsworth.
- Ziman, J., 1968, *Public Knowledge: The Social Dimension of
Science*, Cambridge University Press, Cambridge.