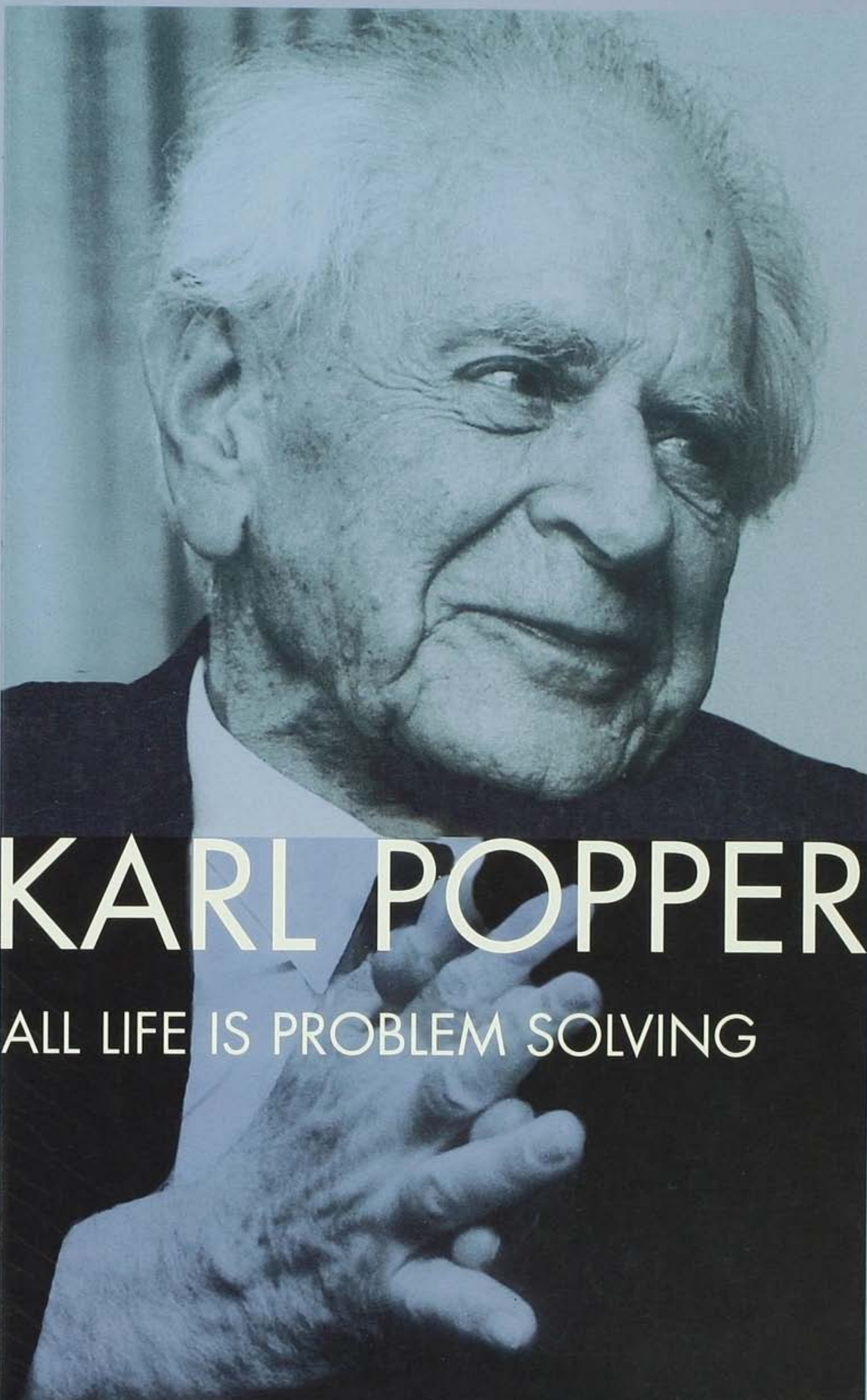
KARL POPPER
ALL LIFE IS PROBLEM SOLVING

ALL LIFE IS
PROBLEM SOLVING

ALL LIFE IS PROBLEM SOLVING

Karl Popper

Translated by Patrick Camiller

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 1999
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Simultaneously published in the USA and Canada
by Routledge
270 Madison Avenue, New York, NY 10016

First published in paperback 2001

Reprinted 2002, 2003, 2005, 2006, 2007

Routledge is an imprint of the Taylor & Francis Group, an informa business

First published in German 1994 by Piper Verlag, Munich

Typeset in Garamond by
M Rules
Printed and bound in Great Britain by
MPG Books Ltd, Bodmin

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
A catalog record for this book has been requested

ISBN 0 415 24992 9

CONTENTS

PUBLISHER'S NOTE

In this translation, four chapters have been omitted from the original German version of *Alles Leben ist Problemlösen*, since all have appeared previously in English, in either identical or slightly varied form, in other works of the author published by Routledge: Chapter 2, 'Wissenschaftliche Reduktion und die essentielle Unvollständigkeit der Wissenschaft', is Addendum 2 of *The Open Universe. An Argument for Indeterminism*, 1982; Chapter 8, 'Über Geschichtsschreibung und über den Sinn der Geschichte', forms the bulk of Chapter 25 of *The Open Society and Its Enemies*, 1945 (5th edition, 1966); and Chapter 10, 'Bemerkungen zur Theorie und Praxis des demokratischen Staates' and Chapter 11, 'Freiheit und intellektuelle Verantwortung', are Chapters 8 and 9 respectively of *The Lesson of This Century*, 1997.

The omitted chapters have been replaced by 'Towards an evolutionary theory of knowledge' (Chapter 5), 'Masaryk and the open society' (Chapter 14), and 'How I became a philosopher without trying' (Chapter 15). All were originally written in English.

'The collapse of communism' (Chapter 12) was also written in English and translated into German for *Alles Leben ist Problemlösen*. This volume publishes the English original for the first time.

This translation by Patrick Camiller has been prepared with the assistance and advice of Melitta Mew and David Miller.

PREFACE

This collection of essays and talks may be seen as a sequel to my book *In Search of a Better World*. Both contain some contributions strongly oriented towards the natural sciences, and others strongly oriented towards history or politics. The title *All Life is Problem Solving* is also the title of Chapter 9, which strongly influenced the short but relevant 'Summary By Way of a Preface' in that earlier collection. Here too, I have tried to give the preface more weight than prefaces usually have.

The selection of chapters has been made with the help and advice of my assistant, Mrs Melitta Mew, and Dr Klaus Stadler from Piper Verlag. I am deeply grateful to them both.

I

The first part of this book is called 'Questions of natural science'. What I have in mind are mainly biology and the unfathomable wealth of living forms.

The more deeply we penetrate into the many areas of biology, from whichever angle, the more unfathomable does the wealth of biological structures prove to be, and the more prodigious their harmonious interplay.

The last chapter in Part I is devoted to Johannes Kepler, the great seeker after harmonies in God's physical creation, and the great discoverer of the three laws that determine the motion of the planets in a highly abstract yet highly harmonious manner. Of the three intellectual giants who together, and with others, created our natural science – the contemporaries Galileo and Kepler, and their successor Newton – Kepler is perhaps the greatest. His is certainly the most attractive, open, and modest personality. All three were passionate seekers and tireless workers; all three toiled extremely hard, often with persistently disappointing results, but they were amply rewarded with the great joy of those who see the world in a new light – differently, more beautifully and harmoniously, and also better than anyone before them – and who then know that their hard work has been crowned with joy, almost undeserved because it could so easily have turned out otherwise.

Of the three great figures, Kepler was the only one who not only worked everything out but honestly and conscientiously wrote it all down. He also understood, as no one else did, that it was the Greek thinkers of the distant past – from Thales to Aristotle, Aristarchus, and Ptolemy – who had bequeathed their boldest ideas to Kepler's inspirational model, Copernicus.

More than in the other two cases, it was his great modesty that again and again helped Kepler to see and learn from his own mistakes – mistakes that could be overcome only with the utmost difficulty. Each of the three intellectual giants was, in his own way, caught up in a superstition. ('Superstition' is a word we should use only with the greatest caution, knowing how little we know and how certain it is that we too, without realizing it, are caught up in various forms of superstition.) Galileo most deeply believed in a natural circular motion – the very belief that Kepler, after lengthy struggles, conquered both in himself and in astronomy. Newton wrote a long book on the traditional (mainly biblical) history of mankind, whose dates he adjusted in accordance with principles quite clearly derived from superstition. And Kepler was not only an astronomer but also an astrologer; he was for this reason dismissed by Galileo and many others.

But Kepler fought against dogmatic forms of his own astrological superstition: he was a self-critical astrologer. He taught that the fate written in the stars was not inexorable but could be mastered by our moral will. It was a major concession to the critics of astrology. Of the three great men, he was perhaps the least dogmatic in his superstition.

II

The second part of this book, 'Thoughts on history and politics', consists of a number of occasional pieces. It offers no advice or prescriptions, least of all infallible ones, but it does make the case for an attitude of responsibility.

I am naturally in favour of democracy – but not in the way that most of its advocates are. Winston Churchill once said: 'Democracy is the worst form of government, only excepting all other forms of government.' We have nothing better than to abide by majority decisions. A majority government is *accountable*, a coalition government much less so, and a minority government still less.

'Democracy' in the sense of 'rule by the people' has practically never existed, and when it has, it has been an arbitrary and unaccountable dictatorship. A government can and should be accountable to the people. Rule by the people cannot be; it is unaccountable.

I am therefore in favour of *democratically elected, constitutional government*, which is quite different from rule by the people. And I am in favour of accountable government – accountable first of all to those who elected it, but also, perhaps still more, morally responsible to humanity.

Never before have there been so many and such dreadful weapons in so many irresponsible hands – a thousand times more than after the two world wars. The fact that this is so and that our political leaders accept it, is something for which they are accountable to us. We must hold them all to blame for it.

Most of our political leaders would be glad to change how things are. But they have inherited from their predecessors a world situation that was constantly deteriorating because of the gang-leaders' arms race; and they seem, however reluctantly, to have come to terms with it. Any interference seems too risky and difficult. So they talk about it as little as possible.

After the wars there was talk – of disarmament. The Western democracies did disarm to quite a considerable degree. But only they did. This was the great idea of the League of Nations and later, after the Second World War, of the United Nations – the idea that the moral and military superiority of these bodies put them under an obligation to keep the peace, until the others had seen and learnt their duty.

No one can doubt that we are on the point of retreating from this position. We do not explain this to the voters: we are afraid of making sacrifices for it. We prefer not to get involved in 'adventures', a word we use to describe our duty.

III

When I try to think about our European and American history, I come to a conclusion rather like that of the historian H.A.L. Fisher, which I have several times quoted before: 'The fact of progress is written plain and large on the page of history; but progress is not a law of nature. The ground gained by one generation may be lost by the next.'

I can and must agree with each of these three statements. But what is the 'progress' of which Fisher rightly says that history informs us, but which is insecure and capable of being lost?

The answer to this question is as clear as it is important. The progress that Fisher has in mind, that we all have in mind, is ethical or moral progress. It is directed towards the peace on earth already promised in the New Testament, when violence of any kind disappears in both the internal and external relations of states. It is progress towards a civilized human society, progress towards the rule of law and a league of all states based upon the rule of law, with the aim of maintaining peace.

This – according to Kant – is our moral duty: the duty of all men of good will; the goal that we must set for history. Since nuclear weapons appeared, it has been a *necessary* goal.

Not only does this goal have good prospects (as the civilized states are at present also the more powerful); it is necessary. The existence of nuclear weapons makes it a necessary goal for everyone willing to stand up for the progress of humanity and civilization. For the only alternative is annihilation.

The goal originally comes from the Augustan age of the Roman Empire and from the New Testament: *Et in terra pax hominibus bonae voluntatis* ('And peace on earth to all men of good will' – or, as it might also be translated, 'And peace on earth to be gained by all men of good will').

From the crime of the First World War (which led to the Kellogg Pact) and from the crime of the Second World War (which led to the United Nations), a sufficiently strong political movement has arisen among all men of good will.

But, as Fisher says, 'the ground gained by one generation may be lost by the next'.

It has been lost. We have to win it back. We must give thought to our duty. And we must remind our politicians that their responsibility does not end with their death (or resignation).

K.R.P.
Kenley, 12 July 1994

Part I

QUESTIONS OF NATURAL SCIENCE

1

THE LOGIC AND EVOLUTION OF SCIENTIFIC THEORY*

The central idea I should like to present in this talk may be expressed in the following way.

The natural as well as the social sciences always start from *problems*, from the fact that something inspires *amazement* in us, as the Greek philosophers used to say. To solve these problems, the sciences use fundamentally the same method that common sense employs, the method of *trial and error*. To be more precise, it is the method of *trying out* solutions to our problem and then discarding the false ones as erroneous. This method assumes that we work with a large number of *experimental* solutions. One solution after another is put to the test and eliminated.

At bottom, this procedure seems to be the only logical one. It is also the procedure that a lower organism, even a single-cell amoeba, uses when trying to solve a problem. In this case we speak of testing movements through which the organism tries to rid itself of a troublesome problem. Higher organisms are able to *learn* through trial and error how a certain problem should be solved. We may say that they too make testing movements – mental testings – and that to learn is essentially to try out one testing movement after another until one is found that solves the problem. We might compare the animal's successful solution to an *expectation* and hence to a *hypothesis* or a *theory*. For the animal's behaviour shows us that it expects (perhaps unconsciously or dispositionally) that in a similar case the same testing movements will again solve the problem in question.

The behaviour of animals, and of plants too, shows that organisms are geared to laws or regularities. They *expect* laws or regularities in

* A talk given on North German Radio (NDR), 7 March 1972.

their surroundings, and I conjecture that most of these expectations are genetically determined – which is to say that they are innate.

A *problem* arises for the animal if an expectation proves to have been wrong. This then leads to testing movements, to attempts to replace the wrong expectation with a new one.

If a higher organism is too often disappointed in its expectations, it caves in. It cannot solve the problem; it perishes.

I would like to present what I have said so far about learning through trial and error in a three-stage model. The model has the following three stages:

1 *the problem;*
2 *the attempted solutions;*
3 *the elimination.*

So, the first stage in our model is *the problem*. The problem arises when some kind of disturbance takes place – a disturbance either of innate expectations or of expectations that have been discovered or learnt through trial and error.

The second stage in our model consists of *attempted solutions* – that is, attempts *to solve the problem*.

The third stage in our model is the *elimination* of unsuccessful solutions.

Pluralism is essential to this three-stage model. The first stage, the problem itself, may appear in the singular; but not the second stage, which I have called 'attempted solutions' in the plural. Already in the case of animals we speak of testing movements, in the plural. There would be little sense in calling one particular movement a testing movement.

Stage 2, the attempted solutions, are thus testing movements and therefore in the plural; they are subject to the *process of elimination* in the third stage of our model.

Stage 3, the *elimination*, is *negative*. The elimination is fundamentally the elimination of *mistakes*. If an unsuccessful or misguided solution is eliminated, the problem remains unsolved and gives rise to new attempted solutions.

But what happens if an attempted solution is eventually successful? Two things happen. First, the successful solution is *learnt*. Among animals this usually means that, when a similar problem appears again, the earlier testing movements, including unsuccessful ones, are briefly and sketchily repeated in their original order; they are run through until the successful solution is reached.

Learning means that unsuccessful or discarded solutions drop more and more to the level of passing references, so that eventually the successful attempt at a solution appears to be almost the only one left. This is the elimination procedure, which depends upon a pluralism of attempted solutions.

The organism may be said to have thus learnt a new *expectation*. We may describe its behaviour by saying that it expects the problem to be solved through testing movements and, in the end, through the final testing movement that is not eliminated.

As we shall soon see, the development of this expectation by the organism has its scientific counterpart in the formation of hypotheses or theories. But before I turn to the formation of scientific theories, I should like to point out another biological application of my three-stage model. My *three-stage model*,

1 *the problem;*
2 *the attempted solutions;*
3 *the elimination,*

may also be understood as the schema of Darwin's theory of evolution. It is applicable not only to the evolution of the individual organism but also to *the evolution of species*. In the language of our three-stage model, a change in either the environmental conditions or the inner structure of the organism produces a *problem*. It is *a problem of species adaptation*: that is, the species can survive only if it solves the problem through a change in its genetic structure. How does this happen in the Darwinian view of things? Our genetic apparatus is such that changes or mutations occur again and again in the genetic structure. Darwinism assumes that, in the terms of our model, these mutations function as Stage 2 *attempted solutions*. Most mutations are fatal: they are deadly for the bearer of the mutation, for the organism in which they occur. But in this way they are *eliminated*, in accordance with Stage 3 of our model. In our three-stage model, then, we must again stress the essential pluralism of the second stage of *attempted solutions*. If there were not *very many* mutations, they would not be worth considering as attempted solutions. We must assume that sufficient *mutability* is essential to the functioning of our genetic apparatus.

Now I can finally turn to my main theme, the theory or logic of science.

My first thesis here is that science is a biological phenomenon. Science has arisen out of prescientific knowledge; it is a quite remarkable continuation of commonsense knowledge, which may in turn be seen as a continuation of animal knowledge.

My second thesis is that our three-stage model is also *applicable to science.*

I mentioned at the outset that, as the Greek philosophers already saw, science starts from *problems*, from *amazement* about something that may be quite ordinary in itself but becomes a problem or a source of amazement for scientific thinkers. My thesis is that each new development in science can be understood only in this way, that its starting point is a *problem* or a *problem situation* (which means the appearance of a problem in a certain state of our accumulated knowledge).

This point is extremely important. The old theory of science taught, and still teaches, that the starting point for science is our sense perception or sensory observation. This sounds at first thoroughly reasonable and persuasive, but it is fundamentally wrong. One can easily show this by stating the thesis: *without a problem, no observation.* If I asked you: 'Please, observe!', then linguistic usage would require you to answer by asking me: 'Yes, but what? *What* am I supposed to observe?' In other words, you ask me to set you a *problem* that can be solved through your observation; and if I do not give you a *problem* but only an *object*, that is already something but it is by no means enough. For instance, if I say to you: 'Please look at your watch', you will still not know what I actually want to have observed. But things are different once I set you the most trivial *problem*. Perhaps you will not be interested in the problem, but at least you will know what you are supposed to find out through your perception or observation. (As an example, you might take the problem of whether the moon is waxing or waning, or which town the book you are reading was published in.)

Why did the old theory wrongly think that in science we start from sense perceptions or observations, and not from problems?

In this respect, the old theory of science was dependent upon the commonsense conception of knowledge. This tells us that our knowledge of the external world is entirely derived from our sense impressions.

I generally have a lot of respect for common sense. I even think that, if we are just a little critical, common sense is the most valuable and reliable counsellor in every possible problem situation. But it is *not always reliable.* And in matters of scientific or epistemological theory, it is extremely important to have a really critical attitude to it.

It is obviously true that our sense organs inform us about the world around us and that they are indispensable for that purpose. But from this we cannot conclude that our knowledge begins with sense perception. On the contrary: our senses, from the point of view of evolutionary theory, are tools that have been formed to solve certain biological *problems*. Apparently, animal and human eyes developed so that living things that are able to change their position and move about may be warned in sufficient time of dangerous encounters with hard objects from which they might receive an injury. From the point of view of evolutionary theory, our sense organs are the outcome of a series of problems and attempted solutions, just as our microscopes or binoculars are. And this shows that, biologically speaking, the problem comes *before* the observation or sense perception: observations or sense perceptions are important aids to our *attempted solutions* and play the main role in their *elimination*. My three-stage model is thus applicable in the following way to the logic or methodology of science.

1 The starting point is always a *problem* or a problem situation.
2 *Attempted solutions* then follow. These always consist of theories, and these theories, being *trials*, are very often wrong: they are and always will be hypotheses or conjectures.
3 In science, too, we learn by *eliminating* our mistakes, by *eliminating* our false theories.

Our three-stage model,

1 *problem;*
2 *attempted solutions;*
3 *elimination,*

may therefore be applied in describing science. This brings us to our central question:

What is distinctive about human science? What is the key difference between an amoeba and a great scientist such as Newton or Einstein?

The answer to this question is that the distinctive feature of science is conscious application of the *critical method*; in Stage 3 of our model, the stage of error elimination, we act in a consciously critical manner.

The critical method alone explains the extraordinarily rapid growth of the scientific form of knowledge, the extraordinary progress of science.

All prescientific knowledge, whether animal or human, is *dogmatic*; and science begins with the invention of the non-dogmatic, critical method.

At any event, the invention of the critical method presupposes a descriptive human *language* in which critical *arguments* can take shape.

Possibly it presupposes even writing. For the essence of the critical method is that our attempted solutions, our theories, and our hypotheses, can be formulated and objectively *presented* in language, so that they become *objects of consciously critical investigation.*

It is very important to appreciate the huge difference between a thought that is only subjectively or privately thought or held to be true, which is a dispositional psychological structure, and the *same* thought when formulated in speech (perhaps also in writing) and thus presented for public discussion.

My thesis is that the step from my unspoken thought: 'It will rain today' to the same spoken proposition 'It will rain today' is a hugely important step, a step over an abyss, so to speak. At first this step, the expression of a thought, does not seem so great at all. But to formulate something in speech means that what used to be part of my personality, my expectations and perhaps fears, is now objectively to hand and therefore available for general critical discussion. The difference is also huge for me personally. The proposition – the prediction, for example – detaches itself from me when it is formulated in speech. It becomes independent of my moods, hopes, and fears. It is *objectified.* It can be *experimentally* endorsed by others as well as by myself, but it can also be *experimentally* disputed. The pros and cons can be weighed and discussed. People can take sides for and against the prediction.

We come here to an important distinction between two meanings of the word *knowledge* [*Wissen*] – *knowledge in the subjective and in the objective sense.* Usually knowledge is thought of as a subjective or mental state. Starting from the verb form 'I know', one explains knowing as a certain kind of belief – that is, a kind of belief that rests upon *sufficient reasons.* This subjective interpretation of the word 'knowledge' has had too strong an influence on the old theory of science. In fact, it is completely useless for a theory of science, because scientific knowledge consists of objective propositions formulated in speech, of hypotheses and problems, not of subjective expectations or convictions.

Science is a product of the human mind, but this product is as objective as a cathedral. When it is said that a proposition is a thought expressed in speech, this is true enough but it does not focus sharply enough on its objectivity. This is connected with an ambiguity of the word 'thought'. As the philosophers Bernard Bolzano and (following him) Gottlob Frege have particularly emphasized, we must distinguish the *subjective thought process* from the *objective content* or the *logical* or *informational* content of a thought. If I say: 'Mohammed's thoughts are very different from Buddha's', I am speaking not of the two men's thought processes but of the logical content of the two doctrines or theories.

Thought processes may stand in causal relationships. If I say: 'Spinoza's theory was *influenced* by Descartes's theory', I am describing a causal relation between two people and stating something about Spinoza's thought processes.

But if I say: 'Spinoza's theory nevertheless contradicts Descartes's on several important points', I am speaking of the objective logical content of the two theories and not about thought processes. The logical content of statements is what I have in mind above anything else when I stress the objective character of human speech. And when I said earlier that only thoughts spoken aloud can be subjected to criticism, I meant that only the logical content of a proposition, not the psychological thought process, can be critically discussed.

I should now like to recall my three-stage model:

1 *problem;*
2 *attempted solutions;*
3 *elimination,*

and my remark that this schema of how new knowledge is acquired is applicable all the way from the amoeba to Einstein.

What is the difference? This question is decisive for the theory of science.

The crucial difference appears in Stage 3, in the *elimination* of attempted solutions.

In the prescientific development of knowledge, *elimination* is something that happens to us: the environment eliminates our attempted solutions; we are not active in the elimination but only passively involved; we *suffer* the elimination, and if it too often destroys our attempted solutions, or if it destroys an attempted solution that was previously successful, it thereby destroys not only the attempted solution but also ourselves as its bearers. This is clear in the case of Darwinian selection.

The crucial novelty of the scientific method and approach is simply that we are actively interested and involved in elimination. The attempted solutions are objectified; we are no longer personally identified with our attempted solutions. However much we may or may not be aware of the three-stage model, the novelty in the scientific approach is that we actively seek to eliminate our attempted solutions. We subject them to criticism, and this criticism operates with every means that we have at our disposal and are able to produce. For example, instead of waiting until our environment refutes a theory or an attempted solution, we try to modify the environment in such a way that it is *as unfavourable as possible* to our attempted solution. We thus put our theories to the test – indeed, we try to put them to the severest test. We do everything to eliminate our theories, for we ourselves should like to discover those theories that are *false*.

The question of how the amoeba differs crucially from Einstein may thus be answered as follows.

The amoeba shuns falsification: its expectation is part of itself, and prescientific bearers of an expectation or a hypothesis are often destroyed by refutation of the expectation or the hypothesis. Einstein, however, made his hypothesis objective. The hypothesis is something outside him, and the scientist can destroy his hypothesis through criticism without perishing along with it. In science we get our hypotheses to die for us.

I have now reached my own hypothesis, the theory that so many supporters of the traditional theory of science have branded as paradoxical. My main thesis is that what distinguishes the scientific approach and method from the prescientific approach is the method of *attempted falsification*. Every attempted solution, every theory, is tested as rigorously as it is possible for us to test it. But a rigorous examination is always an attempt to discover the *weaknesses* in what is being examined. Our testing of theories is also an attempt to detect their weaknesses. The testing of a theory is thus an attempt to refute or to *falsify* the theory.

This does not mean, of course, that a scientist is always happy to falsify one of his own theories. He put the theory forward as an *attempted solution*, and this means that it was meant to stand up to rigorous testing. Many scientists who manage to falsify a promising solution feel personally very disappointed.

Often the aim of falsifying the theory will not be the scientist's personal goal, and quite often a genuine scientist will try to defend against falsification a theory in which he invested great hopes.

This is thoroughly desirable from the point of view of scientific theory; for how else could we distinguish *genuine* falsifications from *illusory* ones? In science we need to form parties, as it were, for and against any theory that is being subjected to serious scrutiny. For we need to have a rational scientific *discussion*, and discussion does not always lead to a clear-cut resolution.

At any event, the *critical* approach is the crucial novelty that makes science what it is, achieved above all through objective, public, linguistic formulation of its theories. This usually leads to a taking of sides and hence to critical discussion. Often the debate is not resolved for many years, as in the case of the famous debate between Albert Einstein and Niels Bohr. Besides, we have no guarantee that every scientific discussion can be resolved. There is nothing that can guarantee scientific progress.

My main thesis, then, is that the novelty of science and scientific method, which distinguishes it from the prescientific approach, is its consciously critical attitude to attempted solutions; it takes an active part in attempts at elimination, in attempts to criticize and falsify.

Conversely, attempts to save a theory from falsification also have their methodological function, as we have already seen. But my thesis is that such a dogmatic attitude is essentially characteristic of *pre*scientific thinking, whereas the critical approach involving conscious attempts at falsification leads to *science* and governs *scientific method.*

Although the taking of sides undoubtedly has a function in scientific method, it is in my view important that the individual researcher should be aware of the underlying significance of attempts at falsification and of sometimes successful falsification. For the scientific method is not *cumulative* (as Bacon of Verulam and Sir James Jeans taught); it is fundamentally *revolutionary*. Scientific progress essentially consists in the replacement of earlier theories by later theories. These new theories must be capable of solving all the problems that the old theories solved, and of solving them at least as well. Thus Einstein's theory solves the problem of planetary motion and macromechanics in general, at least as well as, and *perhaps better than*, Newton's theory does. But the revolutionary theory starts from new assumptions, and in its conclusions it goes beyond and directly contradicts the old theory. This contradiction allows it to devise experiments that can distinguish the old from the new theory, but only in the sense that they can falsify at least one of the two theories. In fact, the experiments may prove the superiority of the surviving theory, but not its truth; and the surviving theory may soon be overtaken in its turn.

Once a scientist has grasped that this is how things stand, he will himself adopt a critical attitude to his own pet theory. He will prefer to test it himself and even to falsify it, rather than leave this to his critics.

One example, of which I am proud, is my old friend the brain scientist and Nobel prizewinner Sir John Eccles. I first met John Eccles when I gave a series of lectures at the University of Otago in Dunedin, New Zealand. For years he had already been experimentally tackling the problem of how the nerve impulse is conveyed through 'synaptic transmission' from one nerve cell to another. A school working mainly in Cambridge around Sir Henry Dale conjectured that molecules of a chemical 'transmitter substance' cross the synapse (which separates nerve cells) and thus convey the stimulus from one cell to another. Eccles's experiments had shown, however, that the transmission took an extremely short time – too short in his view for the transmitter substance – and he therefore developed a theory of purely electrical transmission of both nerve stimuli and nerve inhibitions.

But I may let Eccles speak for himself:

> Until 1945 I held the following conventional ideas about scientific research – first, that hypotheses grow out of the careful and methodical collection of experimental data. This is the inductive idea of science deriving from Bacon and Mill. Most scientists and philosophers still believe that this is the scientific method. Second, that the excellence of a scientist is judged by the reliability of his developed hypotheses, which, no doubt, would need elaboration as more data accumulate, but which, it is hoped, will stand as a firm and secure foundation for further conceptual development. A scientist prefers to talk about the experimental data and to regard the hypothesis just as a kind of working construct. Finally, and this is the important point: it is in the highest degree regrettable and a sign of failure if a scientist espouses an hypothesis which is falsified by new data so that it has to be scrapped altogether.
>
> That was my trouble. I had long espoused an hypothesis which I came to realize was likely to have to be scrapped, and I was extremely depressed about it. I had been involved in a controversy about synapses [. . .], believing in those days that the synaptic transmission between nerve cells was largely electrical. I admitted that there was a late, slow chemical component, but I believed that the fast transmission across the synapse was electrical.
>
> At that time I learnt from Popper that it was not scientifically disgraceful to have one's hypothesis falsified. That was the best news I had had for a long time. I was persuaded by Popper, in fact, to formulate my electrical hypotheses of excitatory and inhibitory synaptic transmission so precisely and rigorously that they invited falsification – and, in fact, that is what happened to them a few years later, very largely by my

> colleagues and myself, when in 1951 we started to do intracellular recording from motoneurones. Thanks to my tutelage by Popper, I was able to accept joyfully this death of the brain-child which I had nurtured for nearly two decades and was immediately able to contribute as much as I could to the chemical transmission story which was the Dale and Loewi brain-child.
>
> I had experienced at last the great liberating power of Popper's teachings on scientific method.
>
> There is a strange sequel. It turned out that I had been too precipitate in my complete rejection of the electrical hypotheses of synaptic transmission. The many types of synapse I had worked on were certainly chemical, but now many electrical synapses are known, and in my book on the synapse [Eccles, 1964][1] there are two chapters on electrical transmission, both excitatory and inhibitory!'[2]

It is worth noting that both Eccles and Dale got it wrong with their pathbreaking theories in neurological research; for both thought that their theories were valid for all synapses. Dale's theory was valid for the synapses on which both were working at the time, but it was no more generally applicable than Eccles's theory. Dale's supporters appear never to have recognized this; they were too sure of their victory over Eccles to realize that both sides were guilty of the same (alleged) sin: namely, 'overhasty generalization without waiting for all the relevant data' (which is, however, never practicable).

Elsewhere, in his Nobel Prize biography, Eccles writes: 'Now I can even rejoice in the falsification of a hypothesis I have cherished as my brain-child, for such falsification is a scientific success.'

This last point is extremely important. We are always learning a whole host of things through falsification. We learn not only *that* a theory is wrong; we learn *why* it is wrong. Above all else, we gain a *new and more sharply focused problem*; and a new problem, as we already know, is the real starting point for a new development in science.

You will perhaps have been surprised that I have so often mentioned my three-stage model. I have done this partly to prepare you for a very similar four-stage model, which is typical of science and the dynamics of scientific development. The four-stage model may be derived from our three-stage model (problem, attempted solutions, elimination), because what we do is call the first stage 'the old problem' and the fourth stage 'the new problems'. If we further replace 'attempted solutions' with 'tentative theories', and 'elimination' with 'attempted elimination through critical discussion', we arrive at the four-stage model characteristic of scientific theory.

So it looks like this:

1 *the old problem;*
2 *formation of tentative theories;*
3 *attempts at elimination* through critical discussion, including experimental testing;
4 *the new problems* that arise from the critical discussion of our theories.

My four-stage model allows a whole series of scientific points to be made.

On the problem. Prescientific and scientific problems are initially practical in nature, but with the four-stage cycle they are soon at least partially replaced by theoretical problems. This means that most new problems arise out of the *criticism of theories*: they are internal to theory. This is already true of the problems in Hesiod's cosmogony, and still more of the problems of the Greek pre-Socratic philosophers; and it is true of most problems in the modern natural sciences. The problems are themselves products of theories, and of the difficulties that critical discussion uncovers in theories. These theoretical problems are fundamentally questions regarding *explanations* or explanatory theories: the tentative answers provided by the theories are in fact *attempted explanations.*

Among practical problems we may include the problems of predicting something. But from the *intellectual* standpoint of *pure science* predictions belong to Stage 3 – to the stage of *critical discussion and examination.* They are intellectually interesting because they allow us, in practice and in the real world, to check the validity of our theories or attempted explanations.

We can also see from our four-stage model that we start in science in the middle of a *cycle* of old problems and end with new problems that function in turn as the starting point for a new *cycle.* Because of the cyclical or periodic character of our model we can start at *any* of the four stages. We can begin with *theories*, at Stage 2 of our model. That is, we can say that the scientist starts from an *old theory* and, by critically discussing and eliminating it, arrives at problems that he then tries to solve through *new theories.* Precisely because of the cyclical character, this interpretation is perfectly satisfactory.

Also in its favour is the fact that we may describe the *creation of* *satisfactory theories* as the *goal* of science. On the other hand, the question of the circumstances in which a theory may be deemed *satisfactory* leads straight back to the *problem as starting point*. For evidently the first demand we make of a theory is that it should *solve* problems in need of explanation, by clearing up the difficulties that constitute the problem.

Finally, we may choose as our starting point the elimination or eradication of theories that have existed up to now. For science may be said always to start from the collapse of a theory. This collapse, this elimination, leads to the problem of replacing the eliminated theory with a better theory.

I personally prefer the *problem* as the starting point, but I am well aware that the *cyclical character of the model* makes it possible to regard any of the stages as the starting point for a new development.

A crucial feature of the new four-stage model is its dynamic character. Each of the stages contains, as it were, an inner logical motivation to go on to the next stage. Science, as it appears in this logical sketch, is a phenomenon to be understood as perpetually *growing*; it is essentially *dynamic*, never something *finished*; there is no point at which it reaches its goal once and for all.

There is also another reason why I prefer the *problem* as starting point. The *distance* between an *old problem* and its successor, the *new problem*, seems to me to be a much more impressive feature of scientific progress than is the distance between old theories and the next generation of new theories that replace the old.

Let us take as an example the Newtonian and Einsteinian theories of gravitation. The distance between the two theories is great, yet it is possible to translate Newtonian theory into Einsteinian language, into the formalism of the so-called tensor calculus. And if one does this – as Professor Peter Havas has done, for example – one finds that the difference between the two theories is merely in the finite speed of gravitational propagation, and thus the finite speed of light c. This means that Havas has succeeded in formulating the Einsteinian theory in such a way that, by replacing the *finite* speed c of gravitational propagation with an *infinite* speed, it becomes identical with the Newtonian theory.

It would be quite misleading to conclude that the whole progress made by the theory lies in the finite speed of gravitational propagation.

I would argue that the progressive and dynamic character of the advance is much clearer if one compares the *problems* discovered by critics of Newtonian theory (Ernst Mach, for instance) with the *problems* discovered by critics of Einsteinian theory (mainly Einstein himself).

Thus, if one compares the old with the new problems, one sees the great distance, the great advance. In effect, only one of the old problems has remained unsolved, the explanation of the so-called Mach principle. This is the requirement that we conceive the *inertia* of heavy masses as an effect of remote cosmic masses. Einstein was very disappointed that his theory did not completely explain this. In fact, his theory of gravity made inertia a result of gravitation; yet if we do away with masses in Einstein's theory of gravity, it collapses into the special theory of relativity, and inertia remains even without being caused by masses.

Einstein himself saw this as one of the main defects of his theory. And the problem of integrating the Mach principle into the theory of gravity has occupied every scientist in this field for half a century.

For reasons such as these, it seems to me better to begin our four-stage model with the *problem*. Anyway the model shows what is new in the dynamic development of science compared with *pre*scientific development – namely, our active involvement in the processes of elimination, through the invention of language, writing, and critical discussion. My main thesis is that science emerged through the invention of critical discussion.

An important conclusion from my main thesis refers to the question of how empirical scientific theories differ from other theories. This is itself not an *empirical* scientific problem but a *theoretical* scientific problem; it is a problem that belongs to the logic or philosophy of science. The answer, which can be derived from my main thesis, is as follows.

An empirical scientific theory differs from other theories because it may be undone by possible experimental results: that is to say, possible experimental results can be described that would falsify the theory if we were actually to obtain them.

I have called the problem of distinguishing empirical scientific theories from other theories the 'demarcation problem' and my proposed solution the 'criterion of demarcation'.

My proposed solution to the demarcation problem is thus the following criterion of demarcation. A theory is part of empirical science if and only if it conflicts with possible experiences and is therefore in principle falsifiable by experience.

I have called this criterion of demarcation the 'falsifiability criterion'.

The falsifiability criterion may be illustrated by many theories.

For example, the theory that vaccination protects against smallpox is falsifiable: if someone who has really been vaccinated still gets smallpox, the theory is falsified.

This example may also be used to show that the falsifiability criterion has problems of its own. If one out of a million vaccinated people gets smallpox, we will hardly consider our theory to be falsified. Rather, we will assume that something was wrong with the vaccination or with the vaccine material. And in principle such an escape route is always possible. When we are faced with a falsification, we can always talk our way out somehow or other; we can introduce an auxiliary hypothesis and reject the falsification. We can '*immunize*' our theories against all possible falsification (to use an expression of Professor Hans Albert's).

It is not always simple, then, to apply the falsifiability criterion. Yet the falsifiability criterion does have its value. It is applicable to the theory of smallpox vaccination even if the application is not always quite so simple. If the proportion of vaccinated people who get smallpox is roughly the same as (or perhaps even greater than) the proportion of unvaccinated people who get smallpox, then all scientists will give up the theory of vaccine protection.

Let us compare this case with that of a theory that in my view is not falsifiable: Freud's theory of psychoanalysis, for instance. Evidently, this theory could in principle be tested only if we could describe some human behaviour that conflicted with the theory. There are such falsifiable theories of behaviour: for example, the theory that a man who has lived a long time and always been honest will not suddenly, if his financial circumstances are secure, become a thief in his old age.

This theory is certainly falsifiable, and I suspect that here and there falsifying instances do occur, so that the theory is simply *false* in the formulation just given.

But in contrast to this theory, there seems to be no conceivable human behaviour that could refute psychoanalysis. If a man saves another's life by risking his own, or if he threatens the life of an old friend – whatever unusual human behaviour we might imagine – it will not be in contradiction with psychoanalysis. In principle, psychoanalysis can always explain the most peculiar human behaviour. It is therefore not empirically falsifiable; it is not testable.

I am not saying that Freud did not have many correct insights. I am arguing that his theory is not empirical science, that it is strictly untestable.

This contrasts with theories such as our vaccination example, but above all with theories in physics, chemistry, and biology.

Since Einstein's theory of gravitation, we have reasons to suppose that Newtonian mechanics is false, even though it is an excellent approximation. Anyway both Newton's and Einstein's theories are falsifiable, although of course it is always possible to talk one's way out of falsification by means of an immunization strategy. Whereas no conceivable human behaviour would contradict Freud's psychoanalysis, a table's behaviour would contradict Newton's theory if it were to start moving around. If the full tea-cup on my table were suddenly to start dancing, to spin, and to turn, it would be a falsification of Newton's theory – especially if the tea did not spill as a result of all the spinning and turning. One may say that mechanics stands in contradiction to a whole host of imaginable behaviour on the part of physical bodies – quite unlike psychoanalysis, which stands in contradiction to no conceivable human behaviour.

Einstein's theory of gravitation would itself be affected by almost any imaginable violation of Newtonian mechanics, precisely because Newtonian mechanics is such a good approximation to Einsteinian mechanics. In addition, however, Einstein looked particularly for cases that, if observed, would refute *his* theory but not Newton's.

Einstein wrote, for example, that if his predicted redshift in the spectrum of the satellites of Sirius and other white dwarfs had not been found, he would have considered his theory as refuted.

It is interesting, moreover, that Einstein himself had an extremely critical attitude to his own theory of gravitation. Although none of the experimental tests (all proposed by himself) proved unfavourable to his theory, he regarded it as not fully satisfactory on theoretical grounds. He was perfectly well aware that his theory, like all theories in natural science, was a *provisional attempt at a solution* and therefore had a *hypothetical* character. But he went into greater detail than that. He gave *reasons* why his own theory should be seen as incomplete, and as inadequate for his own research programme. And he listed a set of requirements that an adequate theory would have to fulfil.

But what he claimed for his original theory of gravitation was that it represented a *better approximation* to the sought-after theory than did Newton's theory of gravitation, and therefore *a better approximation to the truth*.

The idea of *approximation to the truth* is, in my view, one of the most important ideas in the theory of science. It is bound up with the fact that, as we have seen, critical discussion of competing theories is so important in science. But critical discussion is regulated by certain values. It needs a regulative principle or, in Kantian terminology, a regulative idea.

Among the regulative ideas that govern the critical discussion of competing theories, three are of the greatest significance: first, the idea of *truth*; second, the idea of the logical and empirical *content* of a theory; and third, the idea of a theory's *truth content* and of its *approximation to the truth*.

That the idea of *truth* governs critical discussion can be seen from the fact that we critically discuss a theory in the hope of eliminating *false* theories. This shows that we are guided by the idea of looking for *true* theories.

The second regulative idea – the idea of the *content* of a theory – teaches us to look for theories with a high informative content. Tautologies or trivial arithmetical propositions such as $12 \times 12 = 144$ are devoid of content: they do not solve any empirical-scientific problem. Difficult problems can be solved only by theories with a high logical and empirical content.

The size of that content is what may be described as the *boldness* of a theory. The more we assert with a theory, the greater is the risk that the theory will be *false*. So we do seek the truth, but we are interested only in bold, risky truths. Examples of bold theories with a high logical content are once again Newton's and Einstein's theories of gravitation, the quantum theory of atoms, and the theory of the genetic code, which partly solves the problem of heredity.

Bold theories such as these have a high content – that is, a high logical and a high empirical content.

These two concepts of content may be explained as follows. The logical content of a theory is the *class of its consequences*, that is, the set or class of all propositions that can be *logically derived* from the theory in question – which will be the higher, the greater the number of consequences.

Even more interesting, perhaps, is the idea of the *empirical content* of a theory. To understand this idea, let us start from the fact that an *empirical* natural law or an *empirical* theory *rules out* certain observable occurrences. (The theory 'All ravens are black' rules out the existence of white ravens; and observation of a white raven refutes the theory.) But as we have seen, Freudian psychoanalysis does not rule out any observable occurrences. Its logical content is certainly high, but its empirical content is nil.

The empirical content of a theory may thus be described as the set or class of empirical propositions excluded by the theory – which means, however, the set or class of empirical propositions that contradict the theory.

Let us take a very simple illustration. The theory that *there are no* *white ravens* contradicts the statement 'Here is a white raven'. It *rules out*, so to speak, the existence of white ravens. The theory that *all ravens are black* has a far greater empirical content. It rules out not only white ravens but also blue, green, and red ravens: the class of excluded propositions is far greater.

An empirical or observational statement that contradicts a theory may be described as a *possible falsification* or a *potential falsifier* of the theory in question. If a possible falsification is actually observed, then the theory is *empirically falsified*.

The statement 'Here is a white raven' is thus a possible falsification both of the low-content theory that there are no white ravens, and of the high-content theory that all ravens are black.

The statement 'On 10 February 1972 a green raven was delivered to the Zoological Gardens in Hamburg' is a possible falsification or a potential falsifier of the theory that all ravens are black, but also of the theory that all ravens are red or blue. If such a statement, such a potential falsifier, is accepted as true on the basis of observation, then all the theories of which it is a falsifier should be regarded as actually falsified. The interesting thing is that the theory says all the more, the greater the number of its potential falsifiers. It says more, and can clear up more problems. Its *explanatory potential* or its *potential explanatory power* is greater.

From this standpoint, we may once again compare Newton's and Einstein's theories of gravitation. What we find is that the empirical content and the potential explanatory power of Einstein's theory are much greater than those of Newton's. For it asserts much more. It describes not only all kinds of motion that Newton's theory describes, especially the planetary orbits, but also the effect of gravity upon light – a problem area about which Newton had nothing to say either in his theory of gravitation or in his optics. Einstein's theory is thus more risky. It may in principle be falsified by observations that do not touch Newton's theory. The empirical content of Einstein's theory, its quantity of potential falsifiers, is thus considerably greater than the empirical content of Newton's theory. Finally, the potential explanatory power of Einstein's theory is by far the greater. If, for instance, we accept that such optical effects as Einstein's predicted redshift in the spectrum of the satellite of Sirius have been confirmed by observations, then these optical effects are also *explained* by Einstein's theory.

But even if the relevant observations have not yet been made, we can say that Einstein's theory is *potentially* superior to Newton's. It has the greater empirical content and the greater explanatory potential. This means that it is theoretically more interesting. At the same time, however, Einstein's theory is at much greater *risk* than Newton's. It is much more prone to falsification, precisely because the number of its potential falsifiers is greater.

It is therefore much more rigorously testable than Newton's theory, which is itself already very rigorously testable. If Einstein's theory stands up to these tests, if it proves its mettle, we still cannot say that it is true, for it may be falsified in later testing; but we can say that not only its empirical content but also its truth content is greater than those of Newton's theory. This means that the number of *true* statements that can be derived from it is greater than the number that can be derived from Newton's theory. And we can say further that Einstein's theory, in light of the critical discussion that makes full use of results of experimental tests, appears to be a better *approximation to the truth*.

The idea of approximation to the truth – like the idea of truth as a regulative principle – presupposes a *realistic view of the world*. It does *not* presuppose that reality is as our scientific theories describe it; but it does presuppose that there is a reality and that we and our theories – which are ideas we have ourselves created and are therefore always idealizations – can draw closer and closer to an adequate description of reality, if we employ the four-stage method of trial and error. But the method is not enough. We must also be lucky. For the conditions we find here on earth, which make life possible and also the development of human speech, human consciousness, and human science, are extremely rare in the cosmos, even if the cosmos is a long way from being as science describes it. For according to science, the world is almost empty of matter and is mainly filled with chaotic radiation; and in the few places where it is *not* empty, it is filled with *chaotic* matter, usually much too hot for the formation of molecules or too cold for the development of forms of life as we know it. Whether or not there is life elsewhere in the universe, from a cosmological point of view life is an uncommonly rare, quite extraordinary phenomenon. And in the development of life, the critical scientific method is in turn an uncommonly rare development – on any probability calculation, almost infinitely improbable. This means that we hit the jackpot when life and science came into being.

The *realistic view of the world*, together with the idea of approximation to the *truth*, seem to me indispensable for an understanding of the perpetually *idealizing* character of science. Moreover, the realistic view of the world seems to me the only humane world-view: it alone accounts for the fact that there are other people who live, suffer, and die as we do.

Science is a system of the products of human ideas – so far idealism is right. But these ideas are likely to fail when tested against reality. This is why realism is right in the end.

I may be thought for a moment to have gone outside my topic with these remarks on realism and the dispute between realism and idealism. But this is not the case – on the contrary. The realism dispute is highly topical in quantum mechanics, and is thus one of the most up-to-date and most open problems in philosophy of science today.

It will be clear by now that I do not have a neutral attitude to this problem. I am totally on the side of *realism*. But there is an influential idealist school in quantum mechanics. In fact there are all conceivable idealist shades, and one famous quantum physicist has even drawn *solipsistic* consequences from quantum mechanics; he claims that these solipsistic consequences necessarily follow from quantum mechanics.

All I can say in reply is that if this is so, something must be wrong with quantum mechanics, however admirable it may be, and however excellent it is as an approximation to reality. Quantum mechanics has stood up to exceptionally rigorous testing. But from this we can draw conclusions about its proximity to the truth only if we are realists.

The struggle over realism and objectivism in scientific theory will continue for a long time to come. We are dealing here with an open and topical problem. It is also, as has been well said, a problem that in a sense carries scientific theory beyond itself. I hope I have made my own position sufficiently clear on this fundamental problem.

Notes

1 John Eccles, *The Physiology of Synapses*, Springer Verlag, Heidelberg, 1964.
2 John Eccles, *Facing Reality: Adventures of a Brain Scientist*, Springer Verlag, New York, Heidelberg & Berlin, 1970, pp. 105–106.