CONJECTURES

sense similar to that in which processes or things may be said to be parts of the world; that the world consists of facts in a sense in which it may be said to consist of (four dimensional) processes or of (three dimensional) things. They believe that, just as certain nouns are names of things, sentences are names of facts. And they sometimes even believe that sentences are something like pictures of facts, or that they are projections of facts. But all this is mistaken. The fact that there is no elephant in this room is not one of the processes or parts of the world; nor is the fact that a hailstorm in Newfoundland occurred exactly 111 years after a tree collapsed in the New Zealand bush. Facts are something like a common product of language and reality; they are reality pinned down by descriptive statements. They are like abstracts from a book, made in a language which is different from that of the original, and determined not only by the original book but nearly as much by the principles of selection and by other methods of abstracting, and by the means of which the new language disposes. New linguistic means not only help us to describe new kinds of facts; in a way, they even create new kinds of facts. In a certain sense, these facts obviously existed before the new means were created which were indispensable for their description; I say, 'obviously' because a calculation, for example, of the movements of the planet Mercury of 100 years ago, carried out today with the help of the calculus of the theory of relativity, may certainly be a true description of the facts concerned, even though the theory was not yet invented when these facts occurred. But in another sense we might say that these facts do not exist as facts before they are singled out from the continuum of events and pinned down by statements—the theories which describe them. These questions, however, although closely connected with our problem, must be left for another discussion. I have mentioned them only in order to make clear that even should the solutions I have proposed be more or less correct, there would still be open problems left in this field.

1. THE GROWTH OF KNOWLEDGE: THEORIES AND PROBLEMS

MY aim in this lecture is to stress the significance of one particular aspect of science—its need to grow, or, if you like, its need to progress. I do not have in mind here the practical or social significance of this need. What I wish to discuss is rather its intellectual significance. I assert that continued growth is essential to the rational and empirical character of scientific knowledge; that if science ceases to grow it must lose that character. It is the way of its growth which makes science rational and empirical; the way, that is, in which scientists discriminate between available theories and choose the better one or (in the absence of a satisfactory theory) the way they give reasons for rejecting all the available theories, thereby suggesting some of the conditions with which a satisfactory theory should comply.

You will have noticed from this formulation that it is not the accumulation of observations which I have in mind when I speak of the growth of scientific knowledge, but the repeated overthrow of scientific theories and their replacement by better or more satisfactory ones. This, incidentally, is a procedure which might be found worthy of attention even by those who see the most important aspect of the growth of scientific knowledge in new experiments and in new observations. For our critical examination of our theories leads us to attempts to test and to overthrow them; and these lead us further to experiments and observations of a kind which nobody would ever have dreamt of without the stimulus and guidance both of our theories and of our
criticism of them. For indeed, the most interesting experiments and observations were carefully designed by us in order to test our theories, especially our new theories.

In this paper, then, I wish to stress the significance of this aspect of science and to solve some of the problems, old as well as new, which are connected with the notions of scientific progress and of discrimination among competing theories. The new problems I wish to discuss are mainly those connected with the notions of objective truth, and of getting nearer to the truth—concepts which seem to me of great help in analysing the growth of knowledge.

Although I shall confine my discussion to the growth of knowledge in science, my remarks are applicable without much change, I believe, to the growth of pre-scientific knowledge—also—that is to say, to the general way in which men, and even animals, acquire new factual knowledge about the world. The method of learning by trial and error—of learning from our mistakes—seems to be fundamentally the same whether it is practised by lower or by higher animals, by chimpanzees or by men of science. My interest is not merely in the theory of scientific knowledge, but rather in the theory of knowledge in general. Yet the study of the growth of scientific knowledge is, I believe, the most fruitful way of studying the growth of knowledge in general. For the growth of scientific knowledge may be said to be the growth of human knowledge writ large (as I have pointed out in the 1958 Preface to my Logic of Scientific Discovery).

But is there any danger that our need to progress will go unsatisfied, and that the growth of scientific knowledge will come to an end? In particular, is there any danger that the advance of science will come to an end because science has completed its task? I hardly think so, thanks to the infinite nature of our ignorance. Among the real dangers to the progress of science is the likelihood of its being completed, but such things as lack of imagination, sometimes a consequence of lack of real interest, or a misplaced faith in formalization and precision, which will be discussed below in section v; or authoritarianism in one or another of its many forms.

Since I have used the word 'progress' several times, I had better make quite sure, at this point, that I am not mistaken for a believer in a historical law of progress. Indeed I have before now struck various blows against the belief in a law of progress, and I hold that even science is not subject to the operation of anything resembling such a law. The history of science, like the history of all human ideas, is a history of irresponsible dreams, of obstinacy, and of error. But science is one of the very few human activities—perhaps the only one—in which errors are systematically criticized and fairly often, in time, corrected. This is why we can say that, in science, we often learn from our mistakes, and why we can speak clearly and sensibly about making progress there. In most other fields of human endeavour, there is change, but rarely.

---

1 See especially my Poverty of Historicism (2nd edn., 1960), and ch. 16 of the present volume.
CONJECTURES

must be something highly desirable is so deeply ingrained that my trivial
the idea that a high degree of probability (in the sense of the calculus of
probability) must be something highly desirable seems to be so obvious to
content' and of 'relative content'; or in other words, that I should not speak
growth of knowledge, then a high probability (in the sense of the calculus of
reason why we may possibly be our aim as well:

Writing $Ct(a)$ for 'the content of the statement $a$', and $Ct(ab)$ for 'the content of the conjunction $a$ and $b$', we have

$\begin{align*}
(1) & \quad Ct(a) < Ct(ab) > Ct(b) \\
(2) & \quad p(a) > p(ab) < p(b)
\end{align*}$

where the inequality signs of (1) are inverted. Together these two laws, (1) and
(2), state that with increasing content, probability decreases, and vice versa;
or in other words, that content increases with increasing improbability. This
analysis is of course in full agreement with the general idea of the logical
criterion of a statement as the class of all those statements which are logically
entailed by it. We may also say that a statement $a$ is logically stronger than
a statement $b$ if its content is greater than that of $b$—that is to say, if it entails
more than $b$.

This trivial fact has the following inescapable consequences: if growth or
knowledge means that we operate with theories of increasing content, it must
also mean that we operate with theories of decreasing probability (in the
sense of the calculus of probability). Thus if our aim is the advancement or

growth of knowledge, then a high probability (in the sense of the calculus of
probability) cannot possibly be our aim as well: these two aims are in-
compatible.

I found this trivial though fundamental result about thirty years ago, and
have been preaching it ever since. Yet the prejudice that a high probability
must be something highly desirable is so deeply ingrained that my trivial
result is still held by many to be 'paradoxical'.3 Despite this simple result the
idea that a high degree of probability (in the sense of the calculus of
probability) must be something highly desirable seems to be so obvious to
most people that they are not prepared to consider it critically. Dr Bruce
Brooke-Wavell has therefore suggested to me that I should stop talking in
this context of 'probability' and should base my arguments on a 'calculus of
content' and of 'relative content'; or in other words, that I should not speak
about science aiming at improbability, but merely say that it aims at maxi-

mum content. I have given much thought to this suggestion, but I do not
think that it would help: a head-on collision with the widely accepted and

3 See for example J. C. Harsanyi, 'Popper's Improbability Criterion for the Choice of
criterion for the choice of scientific hypotheses: every choice remains a risky guess.
Moreover, the theoretician's choice is the hypothesis most worthy of further critical dis-

218

To avoid these simple results, all kinds of more or less sophisticated
theories have been designed. I believe I have shown that none of them is
successful. But what is more important, they are quite unnecessary. One
merely has to recognize that the property which we cherish in theories and
which we may perhaps call 'verisimilitude'—with a calculus totally
different from the calculus of probability with which it seems to have been
classified.

It should be noted that the problem before us is not a problem of words. I
do not mind what you call 'probability', and I do not mind if you call those
degrees for which the so-called 'calculus of probability' holds by any other
name. I personally think that it is most convenient to reserve the term 'prob-
ability' for whatever may satisfy the well-known rules of this calculus (which
Laplace, Keynes, Jeffreys and many others have formulated, and for which I
have given various formal axiom systems). If (and only if) we accept this
terminology, then there can be no doubt that the absolute probability of a
statement $a$ is simply the degree of its logical weakness, or lack of informative
content, and that the relative probability of a statement $a$, given a statement $b$,
is simply the degree of the relative weakness, or the relative lack of new in-
formative content in statement $a$, assuming that we are already in possession
of the information $b$.

Thus if we aim, in science, at a high informative content—if the growth of
knowledge means that we know more, that we know $a$ and $b$, rather than $a$
alone, and that the content of our theories thus increases—then we have to
admit that we also aim at a low probability, in the sense of the calculus of
probability. And since a low probability means a high probability of being falsified, it
follows that a high degree of falsifiability, or refutability, or testability, is one
of the aims of science—in fact, precisely the same aim as a high informative
content.

The criterion of potential satisfactoriness is thus testability, or improb-
ability: only a highly testable or improbable theory is worth testing, and is

219
CONJECTURES

actually (and not merely potentially) satisfactory if it withstands severe tests—especially those tests to which we could point as crucial for the theory before they were ever undertaken.

It is possible in many cases to compare the severity of tests objectively. It is even possible, if we find it worth while, to define a measure of the severity of tests. (See the Addenda to this volume.) By the same method we can define the explanatory power and the degree of corroboration of a theory.4

IV

The thesis that the criterion here proposed actually dominates the progress of science can easily be illustrated with the help of historical examples. The theories of Kepler and Galileo were unified and superseded by Newton's logically stronger and better testable theory, and similarly Fresnel's and Faraday's by Maxwell's. Newton's theory, and Maxwell's, in their turn, were unified and superseded by Einstein's. In each such case the progress was towards a more informative and therefore logically less probable theory: towards a theory which was more severely testable because it made predictions which, in a purely logical sense, were more easily refutable.

A theory which is not in fact refuted by testing those new and bold and improbable predictions to which it gives rise can be said to be corroborated by these severe tests. I may remind you in this connection of Galileo's discovery of Neptune, of Hertz's discovery of electromagnetic waves, of Eddington's eclipse observations, of Elsasser's interpretation of Davison's maxima as interference fringes of de Broglie waves, and of Powell's observations of the first Yukawa mesons.

All these discoveries represent corroborations by severe tests—by predictions which were highly improbable in the light of our previous knowledge (previous to the theory which was tested and corroborated). Other important discoveries have also been made while testing a theory, though they did not lead to its corroboration but to its refutation. A recent and important case is the refutation of parity. But Lavoisier's classical experiments which show that the volume of air decreases while a candle burns in a closed space, or that the weight of burning iron-filings increases, do not establish the oxygen theory of combustion; yet they tend to refute the phlogiston theory.

Lavoisier's experiments were carefully thought out; but even most so-called 'chance-discoveries' are fundamentally of the same logical structure. For these so-called 'chance-discoveries' are as a rule refutations of theories which were consciously or unconsciously held: they are made when some of our expectations (based upon these theories) are unexpectedly disappointed. Thus the catalytic property of mercury was discovered when it was accidentally found that in its presence a chemical reaction had been speeded up which had not been expected to be influenced by mercury. But neither Oersted's nor Röntgen's nor Becquerel's nor Fleming's discoveries were really accidental, even though they had accidental components: every one of these men was searching for an effect of the kind he found.

We can even say that some discoveries, such as Columbus' discovery of America, corroborate one theory (of the spherical earth) while refuting at the same time another (the theory of the size of the earth, and with it, of the nearest way to India); and that they were chance-discoveries to the extent to which they contradicted all expectations, and were not consciously undertaken as tests of those theories which they refuted.

V

The stress I am laying upon change in scientific knowledge, upon its growth, or its progressiveness, may to some extent be contrasted with the current ideal of science as an axiomatized deductive system. This ideal has been dominant in European epistemology from Euclid's Platonizing cosmology (for this is, I believe, what Euclid's Elements were really intended to be) to that of Newton, and further to the systems of Boscovic, Maxwell, Einstein, Bohr, Schrödinger, and Dirac. It is an epistemology that sees the final task and end of scientific activity in the construction of an axiomatized deductive system.

As opposed to this, I now believe that these most admirable deductive systems should be regarded as stepping stones rather than as ends: as important stages on our way to richer, and better testable, scientific knowledge. Regarded thus as means or stepping stones, they are certainly quite indispensable, for we are bound to develop our theories in the form of deductive systems. This is made unavoidable by the logical strength, by the great informative content, which we have to demand of our theories if they are to be better and better testable. The wealth of their consequences has to be unfolded deductively; for as a rule, a theory cannot be tested except by testing, one by one, some of its more remote consequences; consequences, that is, which cannot immediately be seen upon inspecting it intuitively.

Yet it is not the marvellous deductive unfolding of the system which makes a theory rational or empirical but the fact that we can examine it critically; that is to say, subject it to attempted refutations, including observational tests; and the fact that, in certain cases, a theory may be able to withstand those criticisms and those tests—among them tests under which its predecessors broke down, and sometimes even further and more severe tests. It is in the rational choice of the new theory that the rationality of science lies, rather than in the deductive development of the theory.

Consequently there is little merit in formalizing and elaborating a deductive non-conventional system beyond the requirements of the task of criticizing and testing it, and of comparing it critically with competitors. This critical

4 See especially appendix vi to my *L.Sc.D.*
comparison, though it has, admittedly, some minor conventional and arbitrary aspects, is largely non-conventional, thanks to the criterion of progress. It is this critical procedure which contains both the rational and the empirical elements of science. It contains those choices, those rejections, and those decisions, which show that we have learnt from our mistakes, and thereby added to our scientific knowledge.

Yet perhaps even this picture of science—as a procedure whose rationality consists in the fact that we learn from our mistakes—is not quite good enough. It may still suggest that science progresses from theory to theory and that it consists of a sequence of better and better deductive systems. Yet what I really wish to suggest is that science should be visualized as progressing from problems to problems—to problems of ever increasing depth.

For a scientific theory—an explanatory theory—is, if anything, an attempt to solve a scientific problem, that is to say, a problem concerned or connected with the discovery of an explanation. 6

Admittedly, our expectations, and thus our theories, may precede, historically, even our problems. Yet science starts only with problems. Problems crop up especially when we are disappointed in our expectations, or when our theories involve us in difficulties, in contradictions; and these may arise either within a theory, or between two different theories, or as the result of a clash between our theories and our observations. Moreover, it is only through a problem that we become conscious of holding a theory. It is the problem which challenges us to learn; to advance our knowledge; to experiment; and to observe.

Thus science starts from problems, and not from observations; though observations may give rise to a problem, especially if they are unexpected; that is to say, if they clash with our expectations or theories. The conscious task before the scientist is always the solution of a problem through the construction of a theory which solves the problem; for example, by explaining unexpected and unexplained observations. Yet every worthwhile new theory raises new problems; problems of reconciliation, problems of how to conduct new and previously unthought-of observational tests. And it is mainly through the new problems which it raises that it is fruitful.

Thus we may say that the most lasting contribution to the growth of scientific knowledge that a theory can make are the new problems which it raises, so that we are led back to the view of science and of the growth of knowledge as always starting from, and always ending with, problems—problems of an ever increasing depth, and an ever increasing fertility in suggesting new problems.

---

6 Compare this and the following two paragraphs with my Poverty of Historicism, section 28, pp. 121 ff., and chs. 1 and 16 of this volume.

7 See my L.St.D., especially section 84.
8 Cp. Wittgenstein's Tractatus, especially 4.0141; also 2.161; 2.17; 2.223; 3.11.
9 See especially pp. 56-7 of his remarkable Erkenntnislehre, 2nd edn., 1925.
from Tarski’s analysis how to dodge its inconsistencies; which means, admittedly, the introduction of a certain amount of ‘artificiality’—or caution—into its use.)

Although I may assume in this assembly some familiarity with Tarski’s theory of truth, I may perhaps explain the way in which it can be regarded, from an intuitive point of view, as a simple elucidation of the idea of correspondence with the facts. I shall have to stress this almost trivial point because, in spite of its triviality, it will be crucial for my argument.

The highly intuitive character of Tarski’s ideas seems to become more evident (as I have found in teaching) if we first decide explicitly to take ‘truth’ as a synonym for ‘correspondence with the facts’, and then (forgetting all about ‘truth’) proceed to define the idea of correspondence with the facts.

Thus we shall first consider the following two formulations, each of which states very simply (in a metalanguage) under what conditions a certain assertion (in an object language) corresponds to the facts.

(1) The statement, or the assertion, ‘Snow is white’ corresponds to the facts if, and only if, snow is, indeed, white.

(2) The statement, or the assertion, ‘Grass is red’ corresponds to the facts if, and only if, grass is, indeed, red.

These formulations (in which the word ‘indeed’ is only inserted for ease, and may be omitted) sound, of course, quite trivial. But it was left to Tarski to discover that, in spite of their apparent triviality, they contained the solution of the problem of explaining correspondence with the facts and, with it, truth.

I have said that Schlick’s theory was mistaken, yet I think that certain comments he made (loc. cit.) about his own theory throw some light on Tarski’s. For Schlick says that the problem of truth shared the fate of some others whose solutions were not easily seen because they were mistakenly supposed to lie on a very deep level, while actually they were fairly plain and, at first sight, unimpressive. Tarski’s solution may well appear unimpressive at first sight. Yet its fertility and its power are impressive indeed. This, however, is not my topic here.

Thanks to Tarski’s work, the idea of objective or absolute truth—that is truth as correspondence with the facts—appears to be accepted today with confidence by all who understand it. The difficulties in understanding it seem to have two sources: first, the combination of an extremely simple intuitive idea with a certain amount of complexity in the execution of the technical programme to which it gives rise; secondly, the widespread but mistaken dogma that a satisfactory theory of truth would have to be a theory of true belief—of well-founded, or rational belief. Indeed, the three rivals of the correspondence theory of truth—the coherence theory which mistakes consistency for truth, the evidence theory which mistakes ‘known to be true’ for ‘true’, and the pragmatic or instrumentalist theory which mistakes usefulness for truth—these are all subjective (or ‘epistemic’) theories of truth, in contradistinction to Tarski’s objective (or ‘metalogical’) theory. They are subjective in the sense that they all stem from the fundamental subjectivist position which can conceive of knowledge only as a special kind of mental state, or as a disposition, or as a special kind of belief, characterized, for example, by its history or by its relation to other beliefs.

If we start from our subjective experience of believing, and thus look upon knowledge as a special kind of belief, then we may indeed have to look upon truth—that is, true knowledge—as some even more special kind of belief: as one that is well-founded or justified. This would mean that there should be some more or less effective criterion, if only a partial one, of well-foundedness; some symptom by which to differentiate the experience of a well-founded belief from other experiences of belief. It can be shown that all subjective theories of truth aim at such a criterion: they try to define truth in terms of the sources or origins of our beliefs, or in terms of our operations of verification, or of some set of rules of acceptance, or simply in terms of the quality of our subjective convictions. They all say, more or less, that truth is what we are justified in believing or in accepting, in accordance with certain rules or criteria, of origins or sources of our knowledge, of reliability, or stability, or biological success, or strength of conviction, or inability to think otherwise.

The objective theory of truth leads to a very different attitude. This may be seen from the fact that it allows us to make assertions such as the following: a theory may be true even though nobody believes it, and even though we have no reason for accepting it, or for believing that it is true; and another theory may be false, although we have comparatively good reasons for accepting it.

Clearly, these assertions would appear to be self-contradictory from the point of view of any subjective or epistemic theory of truth. But within the objective theory, they are not only consistent, but quite obviously true.

A similar assertion which the objective correspondence theory would make quite natural is this: even if we hit upon a true theory, we shall as a rule be merely guessing, and it may well be impossible for us to know that it is true.

An assertion like this was made, apparently for the first time, by Xenophanes 11 who lived 2,500 years ago; which shows that the objective theory of truth is very old indeed—antedating Aristotle, who also held it. But only with Tarski’s work has the suspicion been removed that the objective theory of truth as correspondence with the facts may be either self-contradictory (because of the paradox of the liar), or empty (as Ramsey suggested), or barren, or at the very least redundant, in the sense that we can do without it (as I once thought myself).

In my theory of scientific progress I might perhaps do without it, up to a point. Since Tarski, however, I no longer see any reason why I should try to

10 See my Introduction to this volume, ‘On the Sources of Knowledge and of Ignorance’.

11 See the Introduction, p. 26, and ch. 5, p. 152 et, above.
avoid it. And if we wish to elucidate the difference between pure and applied science, between the search for knowledge and the search for power or for powerful instruments, then we cannot do without it. For the difference is that, in the search for knowledge, we are out to find true theories, or at least theories which are nearer than others to the truth—which correspond better to the facts; whereas in the search for theories that are merely powerful instruments for certain purposes, we are, in many cases, quite well served by theories which are known to be false.\(^{12}\)

So one great advantage of the theory of objective or absolute truth is that it allows us to say—with Xenophanes—that we search for truth, but may not know when we have found it; that we have no criterion of truth, but are nevertheless guided by the idea of truth as a \textit{regulative principle} (as Kant or Peirce might have said); and that, though there are no general criteria by which we can recognize truth—except perhaps tautological truth—there are something like criteria of progress towards the truth (as I shall explain presently).

The status of truth in the objective sense, as correspondence to the facts, and its role as a \textit{regulative principle}, may be compared to that of a mountain peak which is permanently, or almost permanently, wrapped in clouds. The climber may not merely have difficulties in getting there—he may not know when he gets there, because he may be unable to distinguish, in the clouds, between the main summit and some subsidiary peak. Yet this does not affect the objective existence of the summit, and if the climber tells us 'I have some doubts whether I reached the actual summit', then he does, by implication, recognize the objective existence of the summit. The very idea of error, or of doubt (in its normal straightforward sense) implies the idea of an objective truth which we may fail to reach.

Though it may be impossible for the climber ever to make sure that he has reached the summit, it will often be easy for him to realize that he has not reached it (or not yet reached it); for example, when he is turned back by an overhanging wall. Similarly, there will be cases when we are quite sure that we have not reached the truth. Thus while coherence, or consistency, is no criterion of truth, simply because even demonstrably consistent systems may be false in fact, incoherence or inconsistency do establish falsity; so, if we are lucky, we may discover inconsistencies and use them to establish the falsity of some of our theories.\(^{13}\)

In 1944, when Tarski published the first English outline of his investigations into the theory of truth (which he had published in Poland in 1933), few philosophers would have dared to make assertions like those of Xenophanes; and it is interesting that the volume in which Tarski's paper was published also contained two subjectivist papers on truth.\(^{14}\)

\(^{12}\) See the discussion of the 'second view' (called 'instrumentalism') in ch. 3, above.


\(^{14}\) See the volume referred to in the preceding note, especially pp. 279 and 336.
There is, however, a similar table in which the epistemological (right hand) side is not based on a mistake.

<table>
<thead>
<tr>
<th>Conjectures</th>
<th>Conjecture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truth</td>
<td>Conjecture</td>
</tr>
<tr>
<td>Testability</td>
<td>Empirical test</td>
</tr>
<tr>
<td>Explanatory or predictive power</td>
<td>Degree of corroboration (that is, report of the results of tests)</td>
</tr>
<tr>
<td>'Verisimilitude'</td>
<td></td>
</tr>
</tbody>
</table>

3. Truth and Content: Verisimilitude versus Probability

Like many other philosophers I am at times inclined to classify philosophers as belonging to two main groups—those with whom I disagree, and those who agree with me. I also call them the verificationists or the justificationist philosophers of knowledge (or of belief), and the falsificationists or fallibilists or critical philosophers of knowledge (or of conjectures). I may mention in passing a third group with whom I also disagree. They may be called the disappointed justificationists—the irrationalists and sceptics.

The members of the first group—the verificationists or justificationists—hold, roughly speaking, that whatever cannot be supported by positive reasons is unworthy of being believed, or even of being taken into serious consideration.

On the other hand, the members of the second group—the falsificationists or fallibilists—say, roughly speaking, that what cannot (at present) in principle be overthrown by criticism is (at present) unworthy of being seriously considered; while what can in principle be so overthrown and yet resists all our critical efforts to do so may quite possibly be false, but is at any rate not unworthy of being seriously considered and perhaps even of being believed—though only tentatively.

Verificationists, I admit, are eager to uphold that most important tradition of rationalism—the fight of reason against superstition and arbitrary authority. For they demand that we should accept a belief only if it can be justified by positive evidence; that is to say, shown to be true, or, at least, to be highly probable. In other words, they demand that we should accept a belief only if it can be verified, or probabilistically confirmed.

Falsificationists (the group of fallibilists to which I belong) believe—as most irrationalists also believe—that they have discovered logical arguments which show that the programme of the first group cannot be carried out: that we can never give positive reasons which justify the belief that a theory is true. But, unlike irrationalists, we falsificationists believe that we have also discovered a way to realize the old ideal of distinguishing rational science from various forms of superstition, in spite of the breakdown of the original inductivist or justificationist programme. We hold that this ideal can be realized, very simply, by recognizing that the rationality of science lies not in

By considering their views about the positive or negative function of argument in science, the first group—the justificationists—may be also nicknamed the 'positivists' and the second—the group to which I belong—the critics or the 'negativists'. These are, of course, mere nicknames. Yet they may perhaps suggest some of the reasons why some people believe that only the positivists or verificationists are seriously interested in truth and in the search for truth, while we, the critics or negativists, are flippant about the search for truth, and addicted to barren and destructive criticism and to the propounding of views which are clearly paradoxical.

This mistaken picture of our views seems to result largely from the adoption of a justificationist programme, and of the mistaken subjectivist approach to truth which I have described.

For the fact is that we too see science as the search for truth, and that, at least since Tarski, we are no longer afraid to say so. Indeed, it is only with respect to this aim, the discovery of truth, that we can say that though we are fallible, we hope to learn from our mistakes. It is only the idea of truth which allows us to speak sensibly of mistakes and of rational criticism, and which makes rational discussion possible—that is to say, critical discussion in search of mistakes with the serious purpose of eliminating as many of these mistakes as we can, in order to get nearer to the truth. Thus the very idea of error—and of fallibility—includes the idea of an objective truth as the standard of which we may fall short. (It is in this sense that the idea of truth is a regulative idea.)

Thus we accept the idea that the task of science is the search for truth, that is, for true theories (even though as Xenophanes pointed out we may never get them, or know them as true if we get them). Yet we also stress that truth is not the only aim of science. We want more than mere truth: what we look for is interesting truth—truth which is hard to come by. And in the natural sciences (as distinct from mathematics) what we look for is truth which has a high degree of explanatory power, in a sense which implies that it is logically improbable truth.

For it is clear, first of all, that we do not merely want truth—we want more truth, and new truth. We are not content with 'twice two equals four', even though it is true: we do not resort to reciting the multiplication table if we are faced with a difficult problem in topology or in physics. More truth is not
enough: what we look for are answers to our problems. The point has been well put by the German humorist and poet Busch, of Max-and-Moritz fame, in a little nursery rhyme—I mean a rhyme for the epistemological nursery: 16

Twice two equals four: 'tis true, 
But too empty, and too trite. 
What I look for is a clue  
To some matters not so light.

Only if it is an answer to a problem—a difficult, a fertile problem, a problem of some depth—does a truth, or a conjecture about the truth, become relevant to science. This is so in pure mathematics, and it is so in the natural sciences. And in the latter, we have something like a logical measure of the depth or significance of the problem in the increase of logical improbability or explanatory power of the proposed new answer, as compared with the best theory or conjecture previously proposed in the field. This logical measure is essentially the same thing which I have described above as the logical criterion of potential satisfactoriness and of progress.

My description of this situation might tempt some people to say that truth does not, after all, play a very big role with us negativists even as a regulative principle. There can be no doubt, they will say, that negativists (like myself) much prefer an attempt to solve an interesting problem by a bold conjecture, even if it soon turns out to be false, to any recital of a sequence of true but uninteresting assertions. Thus it does not seem, after all, as if we negativists had much use for the idea of truth. Our ideas of scientific progress and of attempted problem-solving do not seem very closely related to it.

This, I believe, would give quite a mistaken impression of the attitude of our group. Call us negativists, or what you like: but you should realize that we are as much interested in truth as anybody—for example, as the members of a court of justice. When the judge tells a witness that he should speak 'The truth, the whole truth, and nothing but the truth', then what he looks for is as much of the relevant truth as the witness may be able to offer. A witness who likes to wander off into irrelevancies is unsatisfactory as a witness, even though these irrelevancies may be truisms, and thus part of 'the whole truth'. It is quite obvious that what the judge—or anybody else—wants when he asks for 'the whole truth' is as much interesting and relevant true information as can be got; and many perfectly candid witnesses have failed to disclose some important information simply because they were unaware of its relevance to the case.

Thus when we stress, with Busch, that we are not interested in mere truth but in interesting and relevant truth, then, I contend, we only emphasize a point which everybody accepts. And if we are interested in bold conjectures,

16 From W. Busch, Schelm und Ein (first published posthumously in 1909; p. 28 of the Insel edition, 1952). My attention has been drawn to this rhyme by an essay on Busch as a philosopher which my late friend Julius Kraft contributed to the volume Erziehung und Politik (Essays for Minna Specht, 1960); see p. 262. My translation makes it perhaps more like a nursery rhyme than Busch intended.

17 Similar misgivings are expressed by Quine when he criticizes Peirce for operating with the idea of approaching to truth. See W. V. Quine, Word and Object, New York, 1960, p. 23.
all. Almost at once I found that it was not, and that there was no particular difficulty in applying Tarski's fundamental idea to it.

For there is no reason whatever why we should not say that one theory corresponds better to the facts than another. This simple initial step makes everything clear: there really is no barrier here between what at first sight appeared to be Truth with a capital 'T' and truth in a Tarskian sense.

But can we really speak about better correspondence? Are there such things as degrees of truth? Is it not dangerously misleading to talk as if Tarskian truth were located somewhere in a kind of metrical or at least topological space so that we can sensibly say of two theories—say an earlier theory \( t_1 \) and a later theory \( t_2 \), that \( t_2 \) has superseded \( t_1 \), or progressed beyond \( t_1 \), by approaching more closely to the truth than \( t_1 \)?

I do not think that this kind of talk is at all misleading. On the contrary, I believe that we simply cannot do without something like this idea of a better or worse approximation to truth. For there is no doubt whatever that we can say, and often want to say, of a theory \( t_2 \) that it corresponds better to the facts, or that as far as we know it seems to correspond better to the facts, than another theory \( t_1 \).

I shall give here a somewhat unsystematic list of six types of case in which we should be inclined to say of a theory \( t_1 \) that it is superseded by \( t_2 \) in the sense that \( t_2 \) seems—as far as we know—to correspond better to the facts than \( t_1 \), in some sense or other.

1. \( t_2 \) makes more precise assertions than \( t_1 \), and these more precise assertions stand up to more precise tests.

2. \( t_2 \) takes account of, and explains, more facts than \( t_1 \) (which will include for example the above case that, other things being equal, \( t_2 \)'s assertions are more precise).

3. \( t_2 \) describes, or explains, the facts in more detail than \( t_1 \).

4. \( t_2 \) has passed tests which \( t_1 \) has failed to pass.

5. \( t_2 \) has suggested new experimental tests, not considered before \( t_2 \) was designed (and not suggested by \( t_1 \), and perhaps not even applicable to \( t_1 \)); and \( t_2 \) has passed these tests.

6. \( t_2 \) has unified or connected various hitherto unrelated problems.

If we reflect upon this list, then we can see that the contents of the theories \( t_1 \) and \( t_2 \) play an important role in it. (It will be remembered that the logical content of a statement or a theory \( a \) is the class of all statements which follow logically from \( a \), while I have defined the empirical content of \( a \) as the class of all basic statements which contradict \( a \).) For in our list of six cases, the empirical content of theory \( t_2 \) exceeds that of theory \( t_1 \).

This suggests that we combine here the ideas of truth and of content into one—the idea of a degree of better (or worse) correspondence to truth or of greater (or less) likeness or similarity to truth; or to use a term already mentioned above (in contradistinction to probability) the idea of (degrees of) verisimilitude.

It should be noted that the idea that every statement or theory is not only either true or false but has, independently of its truth value, some degree of verisimilitude, does not give rise to any multi-valued logic—that is, to a logical system with more than two truth values, true and false; though some of the things the defenders of multi-valued logic are hankering after seem to be realized by the theory of verisimilitude (and related theories alluded to in section 3 of the Addenda to this volume).

Once I had seen the problem it did not take me long to get to this point. But strangely enough, it took me a long time to put two and two together, and to proceed from here to a very simple definition of verisimilitude in terms of truth and of content. (We can use either logical or empirical content, and thus obtain two closely related ideas of verisimilitude which however merge into one if we consider here only empirical theories, or empirical aspects of theories.)

Let us consider the content of a statement \( a \); that is, the class of all the logical consequences of \( a \). If \( a \) is true, then this class can consist only of true statements, because truth is always transmitted from a premise to all its conclusions. But if \( a \) is false, then its content will always consist of both true and false conclusions. (Example: 'It always rains on Sundays' is false, but its conclusion that it rained last Sunday happens to be true.) Thus whether a statement is true or false, there may be more truth, or less truth, in what it says, according to whether its content consists of a greater or a lesser number of true statements.

Let us call the class of the true logical consequences of \( a \) the 'truth-content' of \( a \) (a German term 'Wahrheitsgehalt'—reminiscent of the phrase 'there is truth in what you say'—of which 'truth-content' may be said to be a translation, has been intuitively used for a long time); and let us call the class of the false consequences of \( a \)—but only these—the 'falsity-content' of \( a \). (The 'falsity-content' is not, strictly speaking, a 'content', because it does not contain any of the true conclusions of the false statements which form its elements. Yet it is possible—see the Addenda—to define its measure with the help of two contents.) These terms are precisely as objective as the terms 'true' or 'false' and 'content' themselves. Now we can say:

Assuming that the truth-content and the falsity-content of two theories \( t_1 \) and \( t_2 \) are comparable, we can say that \( t_2 \) is more closely similar to the truth, or corresponds better to the facts, than \( t_1 \), if and only if either

(a) the truth-content but not the falsity-content of \( t_2 \) exceeds that of \( t_1 \),

(b) the falsity-content of \( t_2 \) but not its truth-content, exceeds that of \( t_1 \).

If we now work with the (perhaps fictitious) assumption that the content and

---

18 This definition is logically justified by the theorem that, so far as the 'empirical part' of the logical content is concerned, comparison of empirical contents and of logical contents always yield the same results; and it is intuitively justified by the consideration that a statement \( a \) tells the more about our world of experience the more possible experiences it excludes (or forbids). About basic statements see also the Addenda to this volume.
The first point is this. Our idea of approximation to truth, or of verisimilitude, should increase the idea of objective or absolute truth. It is obvious that $Vs(a)$ satisfies our two demands, according to which $Vs(a)$ should increase:

(a) if $CtT(a)$ increases while $CtF(a)$ does not, and
(b) if $CtT(a)$ decreases while $CtF(a)$ does not.

Some further considerations of a slightly technical nature and the definitions of $CtF(a)$ and especially $CtT(a)$ and $Vs(a)$ will be found in the Addenda. Here I want only to discuss three non-technical points.

xii

The first point is this. Our idea of approximation to truth, or of verisimilitude, has the same objective character and the same ideal or regulative character as the idea of objective or absolute truth. It is not an epistemological or an epistemic idea—no more than truth or content. (In Tarski's terminology, it is obviously a 'semantic' idea, like truth, or like logical consequence, and, therefore, content.) Accordingly, we have here again to distinguish between the question 'What do you intend to say if you say that the theory $t_2$ has a higher degree of verisimilitude than the theory $t_1$?' and the question 'How do you know that the theory $t_2$ has a higher degree of verisimilitude than the theory $t_1$?'

We have so far answered only the first of these questions. The answer to the second question depends on it, and it is exactly analogous to the following (absolute rather than comparative) question about truth: 'I do not know—I only guess. But I can examine my guess critically, and if it withstands severe criticism, then this fact may be taken as a good critical reason in favour of it.'

My second point is this. Verisimilitude is so defined that maximum verisimilitude would be achieved only by a theory which is not only true, but completely comprehensible true: if it corresponds to all facts, as it were, and, of course, only to real facts. This is of course a very much more remote and unattainable ideal than a mere correspondence with some facts (as in, say, 'Snow is usually white').

But all this holds only for the maximum degree of verisimilitude, and not for the comparison of theories with respect to their degree of verisimilitude. This comparative use of the idea is its main point; and the idea of a higher or lower degree of verisimilitude seems less remote and more applicable and therefore perhaps more important for the analysis of scientific methods than the—in itself much more fundamental—idea of absolute truth itself.
CONJECTURES

Newton’s dynamics, for example, even though we may regard it as refuted, has of course maintained its superiority over Kepler’s and Galileo’s theories. The reason is its greater content or explanatory power. Newton’s theory continues to explain more facts than did the others; to explain them with greater precision; and to unify the previously unconnected problems of celestial and terrestrial mechanics. The reason for the stability of relative appraisals such as these is quite simple: the logical relation between the theories is of such a character that, first of all, there exist with respect to them those crucial experiments, and these, when carried out, went against Newton’s predecessors. And secondly, it is of such a character that the later refutations of Newton’s theory could not support the older theories: they either did not affect them, or (as with the perihelion motion of Mercury) they could be claimed to refute the predecessors also.

I hope that I have explained the idea of better agreement with the facts, or of degrees of verisimilitude, sufficiently clearly for the purpose of this brief survey.

XIV

A brief remark on the early history of the confusion between verisimilitude and probability may perhaps be appropriate here.

As we have seen, progress in science means progress towards more interesting, less trivial, and therefore less ‘probable’ theories (where ‘probable’ is taken in any sense, such as lack of content, or statistical frequency, that satisfies the calculus of probability) and this means, as a rule, progress towards less familiar and less comfortable or plausible theories. Yet the idea of greater verisimilitude, of a better approximation to the truth, is usually confused, intuitively, with the totally different idea of probability (in its various senses of ‘more likely than not’, ‘more often than not’, ‘seems likely to be true’, ‘sounds plausible’, ‘sounds convincing’). The confusion is a very old one. We have only to remember some of the other words for ‘probable’, such as ‘likely’ which comes originally from ‘like the truth’ or ‘verisimilier’ (‘eoikota’, ‘eikota’, ‘elkos’ etc., in Greek; ‘verisimilis’ in Latin; ‘wahrscheinlich’ in German) in order to see some of the traces, and perhaps some of the sources, of this confusion.

Two at least of the earliest of the Presocratic philosophers used ‘eoikota’ in the sense of ‘like the truth’ or ‘similar to the truth’. Thus we read in Xenophanes (OK, B 35): ‘These things, let us suppose, are like the truth.’

It is fairly clear that verisimilitude or truthlikeness is meant here, rather than probability or degree of incomplete certainty. (Otherwise the words ‘let us suppose’ or ‘let it be conjectured’ or ‘let it be imagined’ would be redundant, and Xenophanes would have written something like, ‘These things, let it be said, are probable.’)

Using the same word (‘eoikota’), Parmenides wrote (OK, B 8, 60):19

236

In this fragment ‘eoikota’ has been most frequently translated as ‘probable’ or ‘plausible’. For example W. Kranz, in Diels-Kranz, Fragmente der Vorsokratiker, 6th edn., translates it ‘wahrscheinlich-identisch’ that is, ‘probable and plausible’; he reads the

19 In this fragment ‘eoikota’ has been most frequently translated as ‘probable’ or ‘plausible’. For example W. Kranz, in Diels-Kranz, Fragmente der Vorsokratiker, 6th edn., translates it ‘wahrscheinlich-identisch’ that is, ‘probable and plausible’; he reads the

10 TRUTH, RATIONALITY, AND THE GROWTH OF KNOWLEDGE

‘Now of this world thus arranged to seem wholly like truth I shall tell you . . .’

Yet already in the same generation or the next, Epicharmus, in a criticism of Xenophanes, seems to have used the word ‘eoikota’ in the sense of ‘plausible’, or something like it (OK, 21 a 15); though the possibility cannot be excluded that he may have used it in the sense of ‘like the truth’, and that it was Aristotle (our source is Met., 1010a) who read it in the sense of ‘plausible’ or ‘likely’. Some three generations later, however, ‘eoikota’ is used quite unambiguously in the sense of ‘likely’ or ‘probable’ (or perhaps even of ‘more frequently than not’) by the sophist Antiphon when he writes (OK, x 60): ‘If one begins a thing well it is likely to end well.’

All this suggests that the confusion between verisimilitude and probability goes back almost to the beginning of Western philosophy: and this is understandable if we consider that Xenophanes stressed the fallibility of our knowledge which he described as uncertain guesswork and at best ‘like the truth’. This phrase, it seems, lent itself to misinterpretation as ‘uncertain and at best of some fair degree of certainty’—that is, ‘probable’.

Xenophanes himself seems to have distinguished clearly between degrees of certainty and degrees of truthlikeness. This emerges from another fragment (quoted above towards the end of chapter 5, p. 153) which says that even if by chance we were to hit upon, and pronounce, the final truth (that is, we may add, perfect truthlikeness), we should not know it. Thus great uncertainty is compatible with greatest truthlikeness.

I suggest that we return to Xenophanes and re-introduce a clear distinction between verisimilitude and probability (using this latter term in the sense laid down by the calculus of probability).

The differentiation between these two ideas is the more important as they have become confused; because both are closely related to the idea of truth, and both introduce the idea of an approach to truth by degrees. Logical probability (we do not discuss here physical probability) represents the idea of approaching logical certainty, or tautological truth, through a gradual diminution of informative content. Verisimilitude, on the other hand, represents the idea of approaching comprehensive truth. It thus combines truth and content while probability combines truth with lack of content.20

The feeling that it is absurd to deny that science aims at probability stems, I suggest, from a misguided ‘intuition’—from the intuitive confusion between the two notions of verisimilitude and of probability which, as it now turns out, are utterly different.

passage thus: ‘This world-arrangement (or world-order) I shall expose to you in all its parts as something probable and plausible.’ In translating ‘(wholly) like truth’ or ‘(wholly) like the truth’, I am somewhat influenced by the line (OK, b 35) quoted above from Xenophanes (and also by K. Reinhardt’s Parmenides, p. 5 f., where Wilamowitz is referred to). See also section viii of the Introduction to the present volume; the quotation from Osiander in section i of ch. 3; section xii of ch. 5, above; and Addendum 6, below.

20 This, incidentally, holds for both, absolute probability, P(a), and relative probability, p(a, b); and there are corresponding absolute and relative concepts of verisimilitude.
People involved in a fruitful critical discussion of a problem often rely, if only unconsciously, upon two things: the acceptance by all parties of the common aim of getting at the truth, or at least nearer to the truth, and a considerable amount of common background knowledge. This does not mean that either of these two things is an indispensable basis of every discussion, or that these two things are themselves 'a priori', and cannot be critically discussed in their turn. It only means that criticism never starts from nothing, even though every one of its starting points may be challenged, one at a time, in the course of the critical debate.

Yet though every one of our assumptions may be challenged, it is quite impracticable to challenge all of them at the same time. Thus all criticism must be piecemeal (as against the holistic view of Duhem and of Quine); which is only another way of saying that the fundamental maxim of every critical discussion is that we should stick to our problem, and that we should subdivide it, if practicable, and try to solve no more than one problem at a time, although we may, of course, always proceed to a subsidiary problem, or replace our problem by a better one.

While discussing a problem we always accept (if only temporarily) all kinds of things as unproblematic: they constitute for the time being, and for the discussion of this particular problem, what I call our background knowledge. Few parts of this background knowledge will appear to us in all contexts as absolutely unproblematic, and any particular part of it may be challenged at any time, especially if we suspect that its uncritical acceptance may be responsible for some of our difficulties. But almost all of the vast amount of background knowledge which we constantly use in any informal discussion will, for practical reasons, necessarily remain unquestioned; and the misguided attempt to question it all—i.e., to start from scratch—can easily lead to the breakdown of a critical debate. (Were we to start the race where Adam started, I know of no reason why we should get any further than Adam did.)

The fact that, as a rule, we are at any given moment taking a vast amount of traditional knowledge for granted (for almost all our knowledge is traditional) creates no difficulty for the falsificationist or fallibilist. For he does not accept this background knowledge; neither is established nor as fairly certain, nor yet as probable. He knows that even its tentative acceptance is risky, and stresses that every bit of it is open to criticism, even though only in a piecemeal way. We can never be certain that we shall challenge the right bit; but since our quest is not for certainty, this does not matter. It will be noticed that this remark contains my answer to Quine's holistic view of empirical tests; a view which Quine formulates (with reference to Duhem), by asserting that our statements about the external world face the tribunal of sense experience.

10 TRUTH, RATIONALITY, AND THE GROWTH OF KNOWLEDGE

not individually but only as a corporate body.21 Now it has to be admitted that we can often test only a large chunk of a theoretical system, and sometimes perhaps only the whole system, and that, in these cases, it is sheer guesswork which of its ingredients should be held responsible for any falsification; a point which I have tried to emphasize—also with reference to Duhem—for a long time past.22 Though this argument may turn a verificationist into a sceptic, it does not affect those who hold that all our theories are guesses anyway.

This shows that the holistic view of tests, even if it were true, would not create a serious difficulty for the fallibilist and falsificationist. On the other hand, it should be said that the holistic argument goes much too far. It is possible in quite a few cases to find which hypothesis is responsible for the refutation; or in other words, which part, or group of hypotheses, was necessary for the derivation of the refuted prediction. The fact that such logical dependencies may be discovered is established by the practice of independence proofs of axiomatized systems; proofs which show that certain axioms of an axiom system cannot be derived from the rest. The more simple of these proofs consist in the construction, or rather in the discovery, of a model—a set of things, relations, operations, or functions—which satisfies all the axioms except the one whose independence is to be shown: for this one axiom—and therefore for the theory as a whole—the model constitutes a counter example.

Now let us say that we have an axiomatized theoretical system, for example of physics, which allows us to predict that certain things do not happen, and that we discover a counter example. There is no reason whatever why this counter example may not be found to satisfy most of our axioms or even all our axioms except one whose independence would be thus established. This shows that the holistic dogma of the 'global' character of all tests or counter examples is untenable. And it explains why, even without axiomatizing our physical theory, we may well have an inkling of what has gone wrong with our system.

This, incidentally, speaks in favour of operating, in physics, with highly analysed theoretical systems—that is, with systems which, even though they may fuse all the hypotheses into one, allow us to separate various groups of hypotheses, each of which may become an object of refutation by counter examples. (An excellent recent example is the rejection, in atomic theory, of the law of parity; another is the rejection of the law of commutation for conjugate variables, prior to their interpretation as matrices, and to the statistical interpretation of these matrices.)

One fact which is characteristic of the situation in which the scientist finds himself is that we constantly add to our background knowledge. If we discard

---
some parts of it, others, closely related to them, will remain. For example, even though we may regard Newton's theory as refuted—that is, his system of ideas, and the formal deductive system which derives from it—we may still assume, as part of our background knowledge, the approximate truth, within limits, of its quantitative formulæ.

The existence of this background knowledge plays an important role in one of the arguments which support (I believe) my thesis that the rational and empirical character of science would vanish if it ceased to progress. I can sketch this argument here only in its barest outline.

A serious empirical test always consists in the attempt to find a refutation, a counter example. In the search for a counter example, we have to use our background knowledge; for we always try to refute first the most risky predictions, the 'most unlikely ... consequences' (as Peirce already saw); which means that we always look in the most probable kinds of places for the most probable kinds of counter examples—most probable in the sense that we should expect to find them in the light of our background knowledge. Now if a theory stands up to many such tests, then, owing to the incorporation of the results of our tests into our background knowledge, there may be, after a time, no places left where (in the light of our new background knowledge) counter examples can with a high probability be expected to occur. But this means that the degree of severity of our test declines. This is also the reason why an often repeated test will no longer be considered as significant or as dangerous.

In these circumstances, a counter example becomes 'most unlikely'.

The idea of simplicity, though intuitively connected with the idea of testability,24 this leads us immediately to our second requirement.

For, secondly, we require that the new theory should be independently testable.25 That is to say, apart from explaining all the explicanda which the new theory was designed to explain, it must have new and testable consequences (preferably consequences of a new kind); it must lead to the prediction of phenomena which have not so far been observed.

This requirement seems to me indispensable since without it our new theory might be ad hoc; for it is always possible to produce a theory to fit any

24 See sections 31–46 of my L.Sc.D. More recently I have stressed (in lectures) the need to relativize comparisons of simplicity to those hypotheses which compete qua solutions of a certain problem, or set of problems. The idea of simplicity, though intuitively connected with the idea of a unified or coherent system or a theory that springs from one intuitive picture of the facts, cannot be analysed in terms of numerical paucity of hypotheses. For every theory can be formulated in one statement; and it seems that, for every theory and every , there is a set of independent axioms (though not necessarily 'organic' axioms in the Warsaw sense).

25 For the idea of an independent test see my paper 'The Aim of Science', Ratio, 1, 1957.
CONJECTURES

given set of explicanda. Thus our two first requirements are needed in order
to restrict the range of our choice among the possible solutions (many of
them uninteresting) of the problem in hand.

If our second requirement is satisfied then our new theory will represent a
potential step forward, whatever the outcome of the new tests may be. For it
will be better testable than the previous theory: the fact that it explains all the
explicanda of the previous theory, and that, in addition, it gives rise to new
tests, suffices to ensure this.

Moreover, the second requirement also ensures that our new theory will, to
some extent, be fruitful as an instrument of exploration. That is to say, it will
suggest to us new experiments, and even if these should at once lead to the
refutation of the theory, our factual knowledge will have grown through the
unexpected results of the new experiments. Moreover, they will confront us
with new problems to be solved by new explanatory theories.

Yet I believe that there must be a third requirement for a good theory. It is
this. We require that the theory should pass some new, and severe, tests.

XIX

Clearly, this requirement is totally different in character from the previous
two. These could be seen to be fulfilled, or not fulfilled, largely by analysing
the old and the new theories logically. (They are ‘formal requirements’.) The
third requirement, on the other hand, can be found to be fulfilled, or not ful-
filled, only by testing the new theory empirically. (It is a ‘material require-
ment’, a requirement of empirical success.)

Moreover, the third requirement clearly cannot be indispensable in the
same sense as are the two previous ones. For these two are indispensable for
deciding whether the theory in question should be at all accepted as a serious
candidate for examination by empirical tests; or in other words, whether it is
an interesting and promising theory. Yet on the other hand, some of the most
interesting and most admirable theories ever conceived were refuted at the
very first test. And why not? The most promising theory may fail if it makes
predictions of a new kind. An example is the marvellous theory of Bohr,
Kramers and Slater of 1924 which, as an intellectual achievement, might
almost rank with Bohr’s quantum theory of the hydrogen atom of 1913. Yet
unfortunately it was almost at once refuted by the facts—by the coincidence
experiments of Bothe and Geiger.27 This shows that not even the greatest
physicist can anticipate the secrets of nature: his inspirations can only be
gueses, and it is no fault of his, or of his theory, if it is refuted. Even Newton’s
theory was in the end refuted; and indeed, we hope that we shall in this way
succeed in refuting, and improving upon, every new theory. And if it is
refuted in the end, why not in the beginning? One might well say that it is
merely a historical accident if a theory is refuted after six months rather than
after six years, or six hundred years.

26 Phil. Mag., 47, 1924, pp. 785 f.

10 TRUTH, RATIONALITY, AND THE GROWTH OF KNOWLEDGE

Refutations have often been regarded as establishing the failure of a
scientist, or at least of his theory. It should be stressed that this is an in-
ductivist error. Every refutation should be regarded as a great success; not
merely a success of the scientist who refuted the theory, but also of the
scientist who created the refuted theory and who thus in the first instance
suggested, if only indirectly, the refuting experiment.

Even if a new theory (such as the theory of Bohr, Kramers, and Slater)
should meet an early death, it should not be forgotten; rather its beauty
should be remembered, and history should record our gratitude to it—for
bequeathing to us new and perhaps still unexplained experimental facts and,
with them, new problems; and for the services it has thus rendered to the
progress of science during its successful but short life.

All this indicates clearly that our third requirement is not indispensable:
even a theory which fails to meet it can make an important contribution to
science. Yet in a different sense, I hold, it is indispensable none the less.
(Bohr, Kramers and Slater rightly aimed at more than making an important
contribution to science.)

In the first place, I contend that further progress in science would become
impossible if we did not reasonably often manage to meet the third require-
ment; thus if the progress of science is to continue, and its rationality not to
decline, we need not only successful refutations, but also positive successes.
We must, that is, manage reasonably often to produce theories that entail
new predictions, especially predictions of new effects, new testable con-
sequences, suggested by the new theory and never thought of before.28 Such
a new prediction was that planets would undergo certain circumstances deviate
from Kepler’s laws; or that light, in spite of its zero mass, would prove to be
subject to gravitational attraction (that is, Einstein’s ‘eclipse-effect’). Another
example is Dirac’s prediction that there will be an anti-particle for every
elementary particle. New predictions of these kinds must not only be pro-
duced, but they must also be reasonably often corroborated by experimental
evidence, I contend, if scientific progress is to continue.

We do need this kind of success; it is not for nothing that the great theories
of science have all meant a new conquest of the unknown, a new success in
predicting what had never been thought of before. We need successes such
as that of Dirac (whose anti-particles have survived the abandonment of
some other parts of his theories), or that of Yukawa’s meson theory. We need
the success, the empirical corroborations, of some of our theories, if only in
order to appreciate the significance of successful and stirring refutations (like
that of parity). It seems to me quite clear that it is only through these tempo-
rary successes of our theories that we can be reasonably successful in attrib-
uting our refutations to definite portions of the theoretical maze. (For we are
reasonably successful in this—a fact which must remain inexplicable for one
who adopts Duhem’s and Quine’s views on the matter.) An unbroken

28 I have drawn attention to ‘new’ predictions of this kind and to their philosophical
significance in ch. 3. See especially pp. 117 f.
sequence of refuted theories would soon leave us bewildered and helpless: we should have no clue about the parts of each of these theories—or of our background knowledge—to which we might, tentatively, attribute the failure of that theory.

Earlier I suggested that science would stagnate, and lose its empirical character, if we should fail to obtain refutations. We can now see that for very similar reasons science would stagnate, and lose its empirical character, if we should fail to obtain verifications of new predictions; that is, if we should only manage to produce theories that satisfy our first two requirements but not the third. For suppose we were to produce an unbroken sequence of explanatory theories each of which would explain all the explicanda in its field, including the experiments which refuted its predecessors; each would also be independently testable by predicted new effects; yet each would be at once refuted when these predictions were put to the test. Thus each would satisfy our first two requirements, but all would fail to satisfy the third.

I assert that, in this case, we should feel that we were producing a sequence of theories which, in spite of their increasing degree of testability, were ad hoc, and that we were not getting any nearer to the truth. And indeed, this feeling may well be justified: this whole sequence of theories might easily be ad hoc. For if it is admitted that a theory may be ad hoc if it is not independently testable by experiments of a new kind but merely explains all the explicanda, then it is clear that the mere fact that the theory is also independently testable cannot as such ensure that it is not ad hoc. This becomes clear if we consider that it is always possible, by a trivial stratagem, to make an ad hoc theory independently testable, if we do not also require that it should pass the independent tests in question: we merely have to connect it (conjunctively) in some way or other with any testable but not yet tested fantastic ad hoc prediction which may occur to us (or to some science fiction writer).

Thus our third requirement, like the second, is needed in order to eliminate ad hoc and other ad hoc theories. But it is needed also for what seem to me even more serious reasons.

I think that we are quite right to expect, and perhaps even to hope, that even our best theories will be superseded and replaced by better ones (though...

29 Dr Jerzy Giedymin (in a paper 'A Generalization of the Refutability Postulate', Studia Logica, 10, 1960, see especially pp. 103 ff.) has formulated a general methodological principle of empiricism which says that our various rules of scientific method must not permit what he calls a 'dilettantist strategy': that is to say, they must exclude the possibility that we shall always win the game played in accordance with these rules: Nature must be able to defeat us at least sometimes. If we drop our third requirement, then we can always win, and need not consider Nature at all, as far as the construction of 'good' theories is concerned: speculations about answers which Nature may give to our questions will play no role in our problem situation which will always be fully determined by our past failures alone.

For our aim as scientists is to discover the truth about our problem; and we must look at our theories as serious attempts to find the truth. If they are not true, they may be, admittedly, important stepping stones towards the truth, instruments for further discoveries. But this does not mean that we can ever be content to look at them as being nothing but stepping stones, nothing but instruments; for this would involve giving up even the view that they are instruments of theoretical discovery; it would commit us to looking upon them as mere instruments for some observational or pragmatic purpose. And this approach would not, I suspect, be very successful, even from a pragmatic point of view: if we are content to look at our theories as mere stepping stones, then most of them will not even be good stepping stones. Thus we ought not to aim at theories which are mere instruments for the exploration of facts, but we ought to try to find genuine explanatory theories: we should make genuine guesses about the structure of the world. In brief, we should not be satisfied with the first two requirements.

Of course, the fulfilment of our third requirement is not in our own hands. No amount of ingenuity can ensure the construction of a successful theory. We also need luck; and we also need a world whose mathematical structure is not so intricate as to make progress impossible. For indeed, if we should cease to progress in the sense of our third requirement—if we should only succeed in refuting our theories but not in obtaining some verifications of predictions of a new kind—we might well decide that our scientific problems have become too difficult for us because the structure (if any) of the world is beyond our powers of comprehension. Even in this case we might proceed, for a time, with theory construction, criticism, and falsification: the rational side of the method of science might, for a time, continue to function. Yet I believe that we should feel that, especially for the functioning of its empirical side, both kinds of successes are essential: success in refuting our theories, and success on the part of some of our theories in resisting at least some of our most determined attempts to refute them.

It may be objected that this is merely good psychological advice about the attitude which scientists ought to adopt—a matter which, after all, is their private affair—and that a theory of scientific method worthy of its name should be able to produce logical or methodological arguments in support of our third requirement. Instead of appealing to the attitude or the psychology of the scientist, our theory of science should even be able to explain his attitude, and his psychology, by an analysis of the logic of the situation in which he finds himself. There is a problem here for our theory of method.

I accept this challenge, and I shall produce three reasons: the first from the idea of truth; the second from the idea of getting nearer to the truth...
(verisimilitude); and the third from our old idea of independent tests and of crucial tests.

(1) The first reason why our third requirement is so important is this. We know that if we had an independently testable theory which was, moreover, true, then it would provide us with successful predictions (and only with successful ones). Successful predictions—though they are not, of course, sufficient conditions for the truth of a theory—are therefore at least necessary conditions for the truth of an independently testable theory. In this sense—and only in this sense—our third requirement may even be said to be ‘necessary’, if we seriously accept truth as a regulative idea.

(2) The second reason is this. If it is our aim to strengthen the verisimilitude of our theories, or to get nearer to the truth, then we should be anxious not only to reduce the falsity content of our theories but also to strengthen their truth content.

Admittedly this may be done in certain cases simply by constructing the new theory in such a way that the refutations of the old theory are explained (saving the phenomena’, in this case the refutations). But there are other cases of scientific progress—cases whose existence shows that this way of increasing the truth content is not the only possible one.

The cases I have in mind are cases in which there was no refutation. Neither Galileo’s nor Kepler’s theories were refuted before Newton: what Newton tried to do was to explain them from more general assumptions, and thus to unify two hitherto unrelated fields of inquiry. The same may be said of many other theories: Ptolemy’s system was not refuted when Copernicus produced his. And though there was, before Einstein, the puzzling experiment of Michelson and Morley, this had been successfully explained by Lorentz and Fitzgerald.

It is in cases like these that crucial experiments become decisively important. We have no reason to regard the new theory as better than the old theory—to believe that it is nearer to the truth—until we have derived from the new theory new predictions which were unobtainable from the old theory (the phases of Venus, the perturbations, the mass-energy equation) and until we have found that these new predictions were successful. For it is only this success which shows that the new theory had true consequences (that is, a truth content) where the old theories had false consequences (that is, a falsity content).

Had the new theory been refuted in any of these crucial experiments then we should have had no reason to abandon the old one in its favour—even if the old theory was not wholly satisfactory. (This was the fate of the Bohr-Kramers-Slater theory.)

In all these important cases we need the new theory in order to find out where the old theory was deficient. Admittedly, the situation is different if the deficiency of the old theory is already known before the new theory is invented; but logically the case has enough similarity with the other cases to regard a new theory which leads to new crucial experiments (Einstein’s mass-energy equation) as superior to one which can only save the known phenomena (Lorentz-Fitzgerald).

(3) The same point—the importance of crucial tests—can be made without appealing to the aim of increasing the verisimilitude of a theory, by using an old argument of mine—the need to make the tests of our explanations independent.10 This need is a result of the growth of knowledge—of the incorporation of what was new and problematic knowledge into background knowledge, with a consequent loss of explanatory power to our theories.

These are my main arguments.

Our third requirement may be divided into two parts: first we require of a good theory that it should be successful in some of its new predictions; secondly we require that it is not refuted too soon—that is, before it has been strikingly successful. Both requirements sound strange. The first because the logical relationship between a theory and any corroborating evidence cannot, it seems, be affected by the question whether the theory is temporarily prior to the evidence. The second because if the theory is doomed to be refuted, its intrinsic value can hardly depend upon delaying the refutation.

Our explanation of this slightly puzzling difficulty is simple enough: the successful new predictions which we require the new theory to produce are identical with the crucial tests which it must pass in order to become sufficiently interesting to be accepted as an advance upon its predecessor, and to be considered worthy of further experimental examination which may eventually lead to its refutation.

Yet the difficulty can hardly be resolved by an inductivist methodology. It is therefore not surprising that inductivists such as John Maynard Keynes have asserted that the value of predictions (in the sense of facts derived from the theory but previously not known) was imaginary; and indeed if the value of a theory would lie merely in its relation to its evidential basis, then it would be logically irrelevant whether the supporting evidence precedes or follows in time the invention of the theory. Similarly the great founders of the hypothetical method used to stress the ‘saving of the phenomena’, that is to say, the demand that the theory should explain known experience. Successful new prediction—of new effects—seems to be a late idea, for obvious reasons; perhaps it was first mentioned by some pragmatist, although the distinction between the prediction of known effects and the prediction of new effects was hardly ever made explicitly. But it seems to me quite indispensable as a part of an epistemology which views science as progressing to better and better explanatory theories, that is, not merely to instruments of exploration, but to genuine explanations.

Keynes’ objection (that it is an historical accident whether this support was known before the theory was proposed, or only afterwards so that it could attain the status of a prediction) overlooks the all-important fact that it is

through our theories that we learn to observe, that is to say, to ask questions which lead to observations and to their interpretations. This is the way our observational knowledge grows. And the questions asked are, as a rule, crucial questions which may lead to answers that decide between competing theories. It is my thesis that it is the growth of our knowledge, our way of choosing between theories, in a certain problem situation, which makes science rational. Now both the idea of the growth of knowledge and that of a problem situation are, at least partly, historical ideas. This explains why another partially historical idea—that of a genuine prediction of evidence (it may be about past facts) not known when the theory was first proposed—may play an important role here, and why the apparently irrelevant time element may become relevant.31

I shall now briefly sum up our results with respect to the epistemologies of the two groups of philosophers I have dealt with, the verificationists and the falsificationists.

While the verificationists or inductivists in vain try to show that scientific beliefs can be justified or, at least, established as probable (and so encourage, by their failure, the retreat into irrationalism), we of the other group have found that we do not even want a highly probable theory. Equating rationality with the critical attitude, we look for theories which, however fallible, progress beyond their predecessors; which means that they can be more severely tested, and stand up to some of the new tests. And while the verificationists laboured in vain to discover valid positive arguments in support of their beliefs, we for our part are satisfied that the rationality of a theory lies in the fact that we choose it because it is better than its predecessors; because it can be put to more severe tests; because it may even have passed them, if we are fortunate; and because it may, therefore, approach nearer to the truth.

APPENDIX: A PRESUMABLY FALSE YET FORMALLY HIGHLY PROBABLE NON-EMPIRICAL STATEMENT

In the text of this chapter I have drawn attention to the criterion of progress and of rationality based on the comparison of degrees of testability or degrees of the empirical content or explanatory power of theories. I did so because these degrees have been little discussed so far.

31 Verificationists may think that the preceding discussion of what I have called here the third requirement quite unnecessarily elaborates what nobody contests. Falsificationists may think otherwise; and personally I feel greatly indebted to Dr Agassi for drawing my attention to the fact that I have previously never explained clearly the distinction between what are called here the second and third requirements. He thus induced me to state it here in some detail. I should mention, however, that he disagrees with me about the third requirement which, as he explained to me, he cannot accept because he can regard it only as a residue of verificationist modes of thought. (See also his paper in the Australian Journal of Philosophy, 39, 1961, where he expresses his disagreement on p. 90.) I admit that there may be a whiff of verificationism here; but this seems to me a case where we have to put up with it, if we do not want a whiff of some form of instrumentalism that takes theories to be mere instruments of exploration.

I always thought that the comparison of these degrees leads to a criterion which is more important and more realistic than the simpler criterion of falsifiability which I proposed at the same time, and which has been widely discussed. But this simpler criterion is also needed. In order to show the need for the falsifiability or testability criterion as a criterion of the empirical character of scientific theories, I will discuss, as an example, a simple, purely existential statement which is formulated in purely empirical terms. I hope this example will also provide a reply to the often repeated criticism that it is perverse to exclude purely existential statements from empirical science and to classify them as metaphysical.

My example consists of the following purely existential theory:

“There exists a finite sequence of Latin elegiac couplets such that, if it is pronounced in an appropriate manner at a certain time and place, this is immediately followed by the appearance of the Devil—that is to say, of a man-like creature with two small horns and one cloven hoof.”

Clearly, this untestable theory is, in principle, verifiable. Though according to my criterion of demarcation it is excluded as non-empirical and non-scientific or, if you like, metaphysical, it is not so excluded by those positivists who consider all well-formed statements and especially all verifiable ones as empirical and scientific.

Some of my positivist friends have indeed assured me that they consider my existential statement about the Devil to be empirical. It is empirical though false, they said. And they indicated that I was mistaking false empirical statements for non-empirical ones.

However, I think that the confusion, if any, is not mine. I too believe that the existential statement is false: but I believe that it is a false metaphysical statement. And why, I ask, should anybody who takes it for empirical think that it is false? Empirically, it is irrefutable. No observation in the world can establish its falsity. There can be no empirical grounds for its falsity.

Moreover, it can be easily shown to be highly probable: like all existential statements, it is in an infinite (or sufficiently large) universe almost logically true, to use an expression of Carnap’s. Thus, if we take it to be empirical, we have no reason to reject it, and every reason to accept it and to believe in it—especially upon a subjective theory of probable belief.

Probability theory tells us even more: it can be easily proved not only that empirical evidence can never refute an almost logically true existential statement, but that it can never reduce its probability.32 (Its probability could be reduced only by some information which is at least ‘almost logically false’, and therefore not by an observational evidence statement.) So the empirical probability or degree of empirical confirmation (in Carnap’s sense) of our statement about the devil-summoning spell must for ever remain equal to unity, whatever the facts may be.

32 This is a consequence of the ‘principle of stability’ of the probability calculus; see theorem (26), section V, of my paper ‘Creative and Non-Creative Definitions in the Calculus of Probability’, Synthese, 15, 1963, No. 2, pp. 167 ff.
CONJECTURES

It would of course be easy enough for me to amend my criterion of demarcation so as to include such purely existential statements among the empirical statements. I merely should have to admit not only testable or falsifiable statements among the empirical ones, but also statements which may, in principle, be empirically 'verified'.

But I believe that it is better not to amend my original falsifiability criterion. For our example shows that, if we do not wish to accept my existential statement about the spell that summons the devil, we must deny its empirical character (notwithstanding the fact that it can easily be formalized in any model language sufficient for the expression of even the most primitive scientific assertions). By denying the empirical character of my existential statement, I make it possible to reject it on grounds other than observational evidence. (See chapter 8, section 2, for a discussion of such grounds; and see chapter 11, especially pp. 275-277, for a discussion and a formalization of a similar argument.)

This shows that it is preferable, as I have been trying to make clear for some considerable time, not to assume uncritically that the terms 'empirical' and 'well-formed' (or 'meaningful') must coincide—and that the situation is hardly improved if we assume, uncritically, that probability or probabilistic 'confirmability' may be used as a criterion of the empirical character of statements or theories. For a non-empirical and presumably false statement may have a high degree of probability, as has been shown here.

REFUTATIONS