

A NEW VIEW OF SCIENTIFIC RATIONALITY

1. — INTRODUCTION

No one can doubt the extraordinary technological power that modern science has given to the western world. As Bertrand Russell put it :

« Science, as a dominant factor in determining the beliefs of educated men, has existed for about 300 years ; as a source of economic technique, for about 150 years. In this brief period it has proved itself an incredibly powerful revolutionary force » ¹.

And even if we set aside its industrial, medical and military applications, and consider science as a purely theoretical system, a systematic attempt to describe and explain the workings of the world we live in, it is still extraordinarily impressive. It is as if, while the rest of mankind had built only mud-huts, a comparatively small number of people in a comparatively short time had built a soaring cathedral.

But when we raise philosophical questions about the cognitive status of the theories of modern science, we get no clear, cogent, positive, and generally agreed answers. Can these theories be known to be true ? There was a time when it was widely believed that they could be. If a theory, such as Newton's, yielded a great variety of striking predictions, all of which were subsequently verified, then surely it *must* be true ? But this confident appraisal suffered two hammer blows, as we shall see. Then can a good scientific theory, though not certainly true, at least be known to have a high probability of being true ? Or if it cannot attain a high probability, can the mass of evidence that bears it out at least *raise* its pro-

1. Bertrand RUSSELL, *The Impact of Science on Society*, London, George Allen and Unwin, 1952, p. 9.

bability ? Although there are philosophers who answer this last question affirmatively, the case for a negative answer is, I believe, overwhelming. Then can we at least know of a successful theory that has superseded a less successful one, that although the former's probability has not been raised and is perhaps stuck at zero, it is closer to the truth than the latter ? Again the answer has to be no. There is at the present time a vociferous group who hate and despise modern science ; and to them these negative results are most welcome ; for they seem to imply that science is just another irrational ideology, on a par with, say, Zande magic. But to those of us who admire science as one of the outstanding achievements of western civilization, these results constitute an urgent challenge.

It is ironic that in the « glad confident morning » of modern science, in the early seventeenth century, when it had made only rather small beginnings, some philosophers were filled with a shining optimism as to the certainty and ultimacy of the theoretical understanding of the world that science could and would achieve in the not too distant future ; whereas today there is widespread philosophical disillusion and scepticism concerning the cognitive status of science, despite the enormous strides it has taken since those early days. In Sections 2 to 4 we will look into the historical reasons for this curiously inverse correlation. In Sections 5 to 7, a way of restoring philosophical confidence in science will be indicated.

2. — THE RISE AND FALL OF THE BACON-DESCARTES IDEAL

In the early seventeenth century there were two great visionary spokesmen for the new science : Francis Bacon and René Descartes. There were large differences between their views of the way in which science should be advanced. But what is striking is the similarities between their visions of what science would achieve. Both men believed that the human understanding, properly regulated, can get to the very bottom of things, unlock Nature's deepest secrets, grasp her ultimate essences. And they both believed that the knowledge to be acquired at this ultimate level could be certain or infallible. We could summarize this « Bacon-Descartes ideal », as I call it, as the claim that science can arrive at *ultimate explanations* that are *certainly true*.

There was a period when it seemed that something approaching the Bacon-Descartes ideal had actually been achieved. When Isaac Newton published his *Principia* in 1687 it did not immediately win general acceptance. There was resistance on the Continent from followers of Descartes who objected to one of its central ideas, and one which Newton himself found perplexing, namely that two bodies, however widely separa-

ted by empty space, act instantaneously upon each other in accordance with the law of universal gravitational attraction ; again, Newton's ideas of absolute space and time ran into sharp opposition from Leibniz, Berkeley and others ; and at the empirical level the erratic ways of the moon gave Newton endless trouble. But as the years passed, empirical difficulties of this latter kind were overcome, one after another ; the theory enjoyed a succession of brilliant predictive successes (concerning, for instance, the perturbations of Saturn and Jupiter when in conjunction, the return of Halley's comet in 1759, which Halley had predicted back in 1682, and the discovery of Neptune, from calculations based on the « misbehaviour » of Uranus, in 1846). As to the theoretical difficulties : the idea of gravitational attraction between bodies at a distance lost its strangeness and began to seem plausible ; indeed, that the attraction must vary inversely with the square of the distance between the bodies now seemed to be dictated by geometrical considerations. Surely this astonishingly successful theory of heavenly and terrestrial mechanics was *true*. And did it not reveal, if not God's entire plan for His creation (for it did not cover electrical, chemical, and biological phenomena), a major part of it ?

3. — HUME'S POINT AND KANT'S ANSWER

And then, in 1739, David Hume published Book I of his *Treatise of Human Nature*. Little notice was taken of it at first ; it lay around for a quarter of a century, like a time-bomb quietly ticking away, before attention was first drawn to the danger, by Thomas Reid in his *Inquiry into the Human Mind*, 1764. Tucked away in Hume's sprawling work are two propositions which, between them, seem to destroy the possibility of genuine scientific *knowledge*. One is that no matter of fact can be settled by reason unaided by experience : anything that we can conceive at all we can conceive as existing, or as not existing ; whether it does in fact exist can be decided only by experience. The other is that neither reason nor experience can, in Hume's words, prove *that those instances, of which we have had no experience, resemble those, of which we have had experience*. If reason and experience together cannot even show that the sun will rise tomorrow, or that the next glass of water a thirsty man drinks will quench his thirst, then they obviously cannot show that a grand scientific theory such as Newton's, with its endless array of predictive implications for the future, is true.

« Answering Hume » has developed into a philosophical industry. After Reid there came Oswald, Beattie, and Priestley, and then the greatest Hume-answerer of all, Immanuel Kant. Kant said of his predeces-

sors : « One cannot observe without feeling a certain pain, how his [Hume's] opponents Reid, Oswald, Beattie and finally Priestley, so entirely missed the point of his problem » ². Kant was entirely persuaded of the truth of Newtonian mechanics. He was also persuaded by Hume that such a soaring theoretical edifice could not be supported merely by a combination of experience and logic. Something more was needed and Kant believed that he could supply this extra something. If we are to understand his answer to Hume we need to bear in mind two important differences between mathematics in his day and in ours. First, non-Euclidean geometries had not yet been developed : Euclidean geometry reigned supreme and was accepted as *the* theory of space, necessarily true and knowable *a priori*. Second, Kant saw pure mathematics as an autonomous system : he would have regarded our twentieth century attempts to reduce mathematics to logic as fundamentally mistaken ; for logic is analytic whereas mathematics, Kant insisted, is synthetic. An analytic truth is empty, has no factual content. But mathematical truths, Kant insisted, are not empty. They have content. However, their content is not empirical ; no experience could conflict with them, nor does their content derive from experience. They have a *pure* content derived not from sensory intuitions (say of coloured surfaces) but from our pure intuitions of space and of time. (Pure intuitions of space supply geometry with its content, while pure intuitions of time provide an idea that is indispensable for arithmetic and the construction of numbers, namely the idea of *succession*.)

Another vital distinction drawn by Kant was of course that between truths that are *a priori* and truths that are *a posteriori*. This distinction concerns the way in which a truth can be known. An *a priori* proposition can be known to be true by reason independently of experience : it is necessarily true. An *a posteriori* truth can be known only from experience, and is only contingently true. Now mathematical truths are not contingent truths known from experience : their truth is necessary and can be known *a priori*. In short, mathematics consists of synthetic *a priori* truths.

But mathematics, though a stronghold of synthetic *a priori* truths, was not, for Kant, their only domain : there is such a thing as Pure Natural Science which likewise consists of synthetic *a priori* truths, of which Kant claimed to have provided a complete inventory and infallible proofs. And armed with these proofs Kant could answer Hume : if, as Hume supposed, our scientific knowledge was constructed only from experience with the help of mere logic, it would collapse ; but it does not collapse because

2. Cf. A.k., IV, p. 258 : Pref. to the *Prolegomena*.

its construction involves, in addition, a steely and immutable framework of synthetic *a priori* principles.

Among the principles which he included in this category were : the permanence of substance, the law of causality, and the principle that all substances coexisting in space are in thoroughgoing reciprocity. These bear an obvious resemblance to certain key principles within classical Newtonian mechanics, namely the conservation of matter, physical determinism, and universal gravitational attraction. Unfortunately for Kant, the rapid and explosive progress of physics since his time has burst asunder that steely framework : the creation of matter and physical indeterminacies are admitted by modern physics, and action-at-a-distance is disallowed. Kant's system is a magnificent ruin and the apriorist answer to Hume is defunct.

4. — THE PROBABILIST ANSWER TO HUME

When it came to be recognized that science cannot achieve certainty, it was widely accepted that it can, however, achieve something approximating certainty, namely high probability. Moritz Schlick was one of many philosophers who took this view : the truth of scientific hypotheses is not absolutely guaranteed, he said ; we must be content if their probability is extremely high ³. *Probabilism*, as we may call this approach to Hume's problem, is at the present time the biggest and busiest branch of the Hume-answering industry. One of its inaugurators was Bolzano, whose *Wissenschaftslehre* was published in 1837. It has been carried on by Jevons, Keynes, Jeffreys, Carnap, and many more recent thinkers.

Bolzano's idea was as follows. If evidence *e* entails a hypothesis *h* then *h* holds in every possible world in which *e* holds and the probability of *h* given *e* is one, or $p(h,e) = 1$. Conversely, if *e* contradicts *h* then *h* holds in no possible world in which *e* holds and $p(h,e) = 0$. If *h* holds in *nearly all* the possible worlds in which *e* holds, then the probability of *h* given *e* is close to one. Again, if *h* holds in *half* the possible worlds in which *e* holds, then $p(h,e) = 1/2$ ⁴.

This idea obviously created the need to find a precise representation of the vague idea of a « possible world ». Carnap met this need with his

3. Moritz SCHLICK, *Allgemeine Erkenntnislehre*, 1918, in *General Theory of Knowledge*, English transl. by Albert E. BLUMBERG, New York/Wien, Springer, 1974, p. 73.

4. *Wissenschaftslehre*, Section 161, in Bernard BOLZANO, *Theory of Science*, transl. by Rolf GEORGE, Oxford, Blackwell, 1972, p. 238 f.

concept of a *state-description* ⁵. Given a suitable and well-specified language, a state-description characterizes every individual in its domain as completely as the language allows. Actually, Carnap chose as the basic units for his probability logic not state-descriptions but structure-descriptions. Two state-descriptions are isomorphic, or have the same structure, if one can be got from the other merely by permuting two or more of the individual constants ; and a structure-description creams off the common content of all those state-descriptions that are isomorphic with one another ⁶. Hintikka chose as the basic units for his system not structure-descriptions but constituents ⁷. A structure-description does not say which individuals instantiate a certain predicate but it does say how many anonymous individuals instantiate it. A constituent does not say how many individuals instantiate it but only that this predicate is instantiated by at least one anonymous individual. It creams off all the qualitative content of a structure-description. An important advantage, from a probabilist point of view, of a constituent over a structure-description is that the initial probability of the former, being independent of the number of individuals in the domain, will remain positive even if the domain is infinite, whereas the initial probability of the latter tends to zero as the number of individuals tends to infinity.

The most influential version of probabilism at the present time is Bayesianism. Let us write $p(h,e)$ to denote the posterior probability of a hypothesis h on evidence e and $p(h)$ to denote the probability of h prior to the arrival of evidence e . The distinctive feature of Bayesianism is, of course, that it is primarily interested not in the absolute level of $p(h,e)$ but in the *ratio* of the value of $p(h,e)$ to that of $p(h)$. Suppose that $p(h,e) = .01$. Then if $p(h) = .0001$ the evidence e will have raised the probability of h a hundredfold and will have strongly confirmed it, even though the absolute level of $p(h,e)$ is low. Conversely, if $p(h) = .1$ and $p(h,e) = .01$, then e will have disconfirmed h . If $p(h,e) = p(h)$ then e is neutral to h .

However Bayesianism does not avoid some severe difficulties which confront any version of probabilism. One of these is connected with the fact that if the prior probability of h is zero, then no evidence, however seemingly favourable to h , can raise its posterior probability above zero. Now the laws of physics are precise and universal : they apply to all space-time regions. And in nearly all systems of probability logic (Hin-

5. Rudolf CARNAP, *Logical Foundations of Probability*, London, Routledge and Kegan Paul, 1950, p. 70 f.

6. *Ibid.*, p. 116 f.

7. Jaakko HINTIKKA, *Logic, Language-Games and Information*, Oxford, Clarendon Press, 1973, chap. I and XI, and *Knowledge and the Known*, Boston, Reidel, 1974, chap. 7.

tikka's is an exception) the probability of such a statement is always zero. Various attempts have been made to get round this major difficulty. Jeffreys tried to overcome it with the help of his simplicity-postulate⁸. After providing an admirably clearcut definition of simplicity, he postulated that the simplest of a set of alternative hypotheses all compatible with the given evidence is, other things being equal, the one that is the most probable. But it turned out that his simplicity-postulate leads to inconsistencies, as Popper pointed out⁹. The underlying reason for this is that the simpler, in Jeffreys's sense, a hypothesis is, the stronger it is or the more it asserts ; but it is a paramount principle of probability logic that if one hypothesis says *more* than another then, other things being equal, it has a *lower* probability than the other (or at any rate not a higher probability ; the probability of both hypotheses may be zero).

Hintikka developed a system in which universal laws can have a positive probability ; for the probability of a law is the sum of the probabilities of each of the constituents that entail it ; and these constituents can have positive probabilities, for the reason indicated above. And if a sufficient amount of favourable evidence comes in, the probability of a law can rise to a high level. But his system involves something analogous to Jeffreys's simplicity-postulate : a time comes when it is the *strongest* of all the unrefuted constituents that is treated as the most probable, contrary to the paramount principle mentioned above. It should be the weakest unrefuted constituent that has the highest probability. And as evidence comes in, a higher proportion of the probabilities of the refuted constituents should accrue to this weakest constituent than to unrefuted constituents that are stronger than it. Now this constituent is the one that asserts *every existential statement* expressible in the language, and hence *denies* every law-statement.

Another difficulty for probabilism, which I consider insuperable, is the following. Let E be some considerable body of evidence and suppose that $p(h, E)$ is much higher than $p(h)$. But before we can conclude that h is very well confirmed we need some assurances concerning E . One is that E does not contain misinformation : all the evidence-statements comprised in it should be verified. And many philosophers (myself included) hold that that is not possible. However that is not the difficulty to which I now wish to draw attention. Suppose for the sake of the argument that E is known to be true. We still need a further assurance. For suppose that there were a bit of evidence in our possession that is not

8. Harold JEFFREYS, *Theory of Probability*, Oxford, Clarendon Press, 2d ed., 1948, p. 100 f.

9. Karl R. POPPER, *The Logic of Scientific Discovery*, London, Hutchinson, 1959, Appendix * viii.

included in E and which actually refutes h . Then the fact that $p(h, E)$ is high (or higher than $p(h)$) is unimportant : the important fact is that the probability of h on *all* the evidence in our possession is zero. In other words, we need to know that E contains (i) *only* evidence-statements that are known by us at the present time to be true and (ii) *all* such statements. This is the famous Requirement of Total Evidence. Unless we can be assured that E satisfies this requirement, we are not entitled to base an appraisal of h on the mere fact that $p(h, E)$ is high.

But now consider whether a person X could ever know, of a given E , that it satisfies this requirement. To make X 's task easier, assume (counter-factually) that, confronted by a particular evidence-statement, X can tell unhesitatingly whether or not he knows it to be true, and that he checks through every evidence-statement contained in E and finds in each case that he does indeed know it to be true. That establishes (i) that E contains *only*, but not (ii) that it contains *all*, evidence-statements known to him. It is as if, to vary Descartes's simile, he needed to establish that a certain basket of apples contains (i) none that are rotten and (ii) all that are not rotten. To verify that E satisfies the total evidence requirement X would need to check that everything outside E is *not* evidence known to him. And the situation becomes even worse if the requirement is modified to one of total *relevant* evidence. To verify the claim that E satisfies this latter requirement, he would need to check that everything inside E is both known to him and relevant to the hypothesis in question, and *also* that everything outside E is not known to him or, if known, is not relevant to the hypothesis. And it is obvious that even if the first part of this task could be completed, the second part could not.

5. — POPPER : CORROBORATION AND VERISIMILITUDE

Hume's problem could be restated as follows : inductive inferences (that is, inferences from observed to unobserved instances) are (1) *indispensable* (both in everyday life and in science) and (2) *unjustifiable* (neither reason nor experience can provide any justification for them). Hume himself recommended « carelessness and inattention » as the *only* remedy for this painful dilemma¹⁰. We might say that on this view (a softer version of which has been taken up by Strawson and Ayer in more recent times) the « rationality » both of common sense and of science lies in the fact that the human mind operates in a robustly non-logical way.

10. Cf. *A Treatise of Human Nature*, end of Book I, Part IV, Section II.

So Hume retained both (1) and (2), despite their mutual antagonism. Both Kant and the probabilists retained (1) and rejected (2). Kant rejected (2) on the ground that logic and experience are reinforced by a framework of synthetic *a priori* principles ; probabilists rejected (2) on the ground that classical deductive logic can be generalized and extended to probability logic within which inductive inferences can be accommodated. As we have seen, these two answers fail ; and Hume's « answer » (carelessness and inattention) is really no answer. A remaining possibility is to retain (2) and reject (1). This is what Karl Popper did ¹¹. According to him there are, in science, no inferences from evidence to laws and theories. The latter remain forever conjectural hypotheses, and all scientific inferences are deductive. From these conjectural theories are deductively derived various predictive consequences ; and the latter, especially the more novel and unexpected ones among them, are put to the test. If a test-result is negative there is a deductive inference from this to the falsity of the theory under test. If a theory has stood up to severe testing it is said to be highly *corroborated* ; but this does not mean that it is very probably true or anything of that sort ; it means only that the theory has performed very well so far. The chief rule of Popper's methodology is that we should provisionally adopt, as the best theory in its field, the one that is best corroborated.

An important advantage of this rule, compared with one based on probabilities, is that it is not prejudiced against *powerful* scientific theories. On the contrary, the more powerful a theory is the more corroborable it is ; and the most corroborable of a set of competing theories may go on to become the best corroborated of them, if it survives severe testing. But the question remains as to why, from Popper's anti-inductivist point of view, which prohibits inferences from evidence to the probable truth of a theory, we should regard the best *corroborated* of a set of competing theories as being, at present, the *best* theory in the set. Were it the sole unfalsified survivor in the set, we would have a good reason to prefer it. But there are bound to be other unrefuted hypotheses : for instance, weaker ones that are strictly entailed by the best corroborated theory. These will be, typically, less corroborable and less well corroborated than it. They will also be « safer » in the sense that they are less likely to be overthrown by future tests. Why should we prefer it to them ?

Popper's main answer has been that if one theory is better corroborated than another, then « we will, in general, have *reason to believe* that

11. He first published this idea in 1933 : cf. K. R. POPPER, *op. cit. supra* n. 9, Appendix * i.

the first is a better approximation to the truth than the second »¹². In other words, we have reason to believe that the first has the greater verisimilitude. But this answer involves a lurch into inductivism. Consider two hypotheses, *A* and *B*, which diverge but which have equal amounts of content, in that the consequences of *A* are in one:one correspondence with those of *B*. (I call such hypotheses « incongruent counterparts ».) Then the verisimilitude-appraisal that *A* is closer to the truth than *B* carries the implication that, if a crucial experiment between them, of a hitherto untried kind, is carried out in the future, then *A* is more likely to pass this test than is *B*. For we can think of the empirical consequences of *A* and *B* on the analogy of two urns containing equal numbers of balls, some white (= true), the others black (= false), and with urn *A* containing a higher proportion of white balls ; and a crucial experiment, that is, an experiment directed at certain divergent consequences of *A* and *B*, is analogous to selecting equal numbers of balls from the two urns. It is more probable that those selected from urn *A* will be white than that those selected from *B* will be. Now a corroboration-appraisal, Popper has insisted, is « an evaluating *report of past performance*... [and] *says nothing whatever about future performance* »¹³. But verisimilitude-appraisals, as we have just seen, *do* carry predictive implications as to future performance. Hence if corroboration-appraisals are taken to provide some sort of justification for verisimilitude-appraisals they *will* be saying something indirectly, or at one remove, about future performance.

6. — THE OPTIMUM AIM FOR SCIENCE

Our conclusions so far have all been negative. Philosophical confidence in science will hardly be restored if it essentially involves invalid inferences *a la* Hume, Ayer¹⁴ and Strawson¹⁵. The Kantian idea of stiffening science with a framework of synthetic *a priori* principles has collapsed.

12. *Realism and the Aim of Science*, ed. by William Warren BARTLEY, London, Hutchinson, 1983, III, p. 58.

13. *Objective Knowledge*, Oxford, Clarendon Press, 1972, p. 18.

14. Alfred J. AYER, in *The Problem of Knowledge*, Harmondsworth, Penguin Books, 1956, p. 80, said that there are certain unbridgeable logical gaps that we are simply to take « in our stride ». He did not say when a gap becomes *too* wide to be taken in our stride.

15. Peter F. STRAWSON, in *Introduction to Logical Theory*, London, Methuen, 1952, said that evidence may *conclusively* establish a theory even though it does not entail it (p. 234). He did not spell out what conditions must be satisfied for this important relation of non-deductive proof to hold. On the contrary, he insisted that no such conditions can be specified (p. 248). He adopted a « sealed lips » policy : *there exist* valid non-deductive inferences but we *cannot say* what their nature is.

The probabilist idea that it is rational for a person X to prefer, given a set of competing hypotheses, the one that is best confirmed, in some probabilist sense, by the total evidence turns out to be unworkable. And the justification for Popper's idea that it is rational to prefer the one that is best corroborated, in his non-probabilist and non-inductivist sense, turns out to have a tacitly inductivist character : there is a slide from, in Humean language, instances of which we have had experience (past performance) to expectations about instances of which we have had no experience (future performance).

Then has the idea of scientific rationality broken down ? Are there never any good cognitive reasons for accepting a scientific theory as the best one in its field at the present time ? It is to this problem that my *Science and Scepticism*¹⁶ is addressed. This book is in the Popperian tradition, but it departs from his views in various ways. The idea of one theory being closer to the truth, or having more verisimilitude, than another plays no role in it. A new theory of the rational, or at least quasi-rational, acceptance of statements into the empirical basis is provided. (Briefly, the idea is that a perceptual experience is treated, not as a premise from which a basic statement about the external world is inferred, but as an explanandum for which a conjectural explanans has been found in which the basic statement figures essentially.) Another difference concerns the comparative measure of the testable content of two theories. It turned out that none of Popper's measures could handle the case where theory A goes beyond *and revises* theory B , a case which we very much need to be able to handle, since theories that satisfy Bohr's famous Principle of Correspondence are of that kind. I dealt with this difficulty by generalizing Popper's original measure, which relied on the subclass relation between the theories' classes of potential falsifiers, with the help of the idea, touched on above, of *incongruent counterparts*. Since their consequences are in one:one correspondence, incongruent counterparts have equal amounts of testable content ; and I say that A has more testable content than B if there exists a counterpart B' of B such that every potential falsifier of B' is a potential falsifier of A but not vice versa. (This allows that B' may be a *congruent* counterpart of B , or logically equivalent to B .)

Concerning the central problem of the rational acceptance of scientific hypotheses, my idea was this. Suppose that we could establish that a certain aim is *the optimum aim* for science ; and suppose, further, that we could establish that the best corroborated theory in its field, if there

16. Princeton, University Press/London, Hutchinson, 1984.

is just one, is the one, at the present time, that best fulfils this aim. Then we would have the best possible reason to accept this theory.

But how can we possibly arrive at the optimum aim for science ? I began by laying down certain obvious adequacy requirements for any aim for science. The two most important ones are that the aim must be *coherent* (it must not contain diverse components that pull in opposite directions) and *feasible* (it must be capable of being fulfilled). The question then becomes : is there a maximumly ambitious aim that conforms with these adequacy-requirements and which would no longer do so if it were further strengthened ? If there is, then this aim would *dominate* any other aim that conforms with them, since it would contain an aspiration not contained in the other one but not vice versa.

My starting point in the search for such a dominating aim was what I call the Bacon-Descartes ideal, which I interpret as saying that it is the goal of science to render all phenomena explainable and predictable by rigorous deduction from (appropriate initial conditions and) universal principles that are : certainly true, ultimate, unified, and exact. It seems obvious that, as a first step in attempting to turn this into a feasible ideal, it should be recast as the aim of *progressing towards*, rather than actually attaining, these far-off and perhaps unattainable goals. Thus « *progressivised* », the aim of science becomes to advance with explanatory theories that are ever more :

- (A) probable ;
- (B1) deep ;
- (B2) unified ;
- (B3) predictively powerful ;
- (B4) exact.

I call the (A)-component of this version of the Bacon-Descartes ideal its security-pole and the (B)-components its depth-pole. As we shall see, it is not, as it stands, a coherent aim of science because its two poles pull in opposite directions. Let us start by considering its depth-pole.

I will not reproduce here the rather technical elucidations given in my book of the ideas of one theory being deeper, and more unified than another, though I will say something later about the main idea behind them. Actually, it turns out that the conditions for one theory to be deeper are essentially the same as those for it to be more unified than another. Instead of two distinct components, (B1) and (B2), we really have one component, call it (B1 - 2). As to (B3), the demand for greater predictive power, it obviously calls for excess testable content. And the same is true of (B4), or the demand for increasing exactitude. Actually, (B3) and (B4) are not really two distinct components, but merge into one component, call it (B3 - 4), just as (B1) and (B2) merged into (B1 - 2). For a univer-

sal theory can always be restated in such a way that it ascribes *one* predicate (which may be very complex) to *everything* ; and we can say that one empirical theory is more exact than another if, when so restated, the predicate it ascribes is more exact than the one the other ascribes to everything. But in that case the former theory will be more testable and have greater predictive power.

Thus all the (B)-components of our progressivised version of the Bacon-Descartes ideal are calling, directly or indirectly, for increasing testable content as science progresses. Now consider component (A). A fundamental negative result obtained by Popper¹⁷ was that the probability of hypotheses, other things being equal, tends to vary *inversely* with their testability. (I inserted the words « tends to » to allow for the possibility that two hypotheses, one of which is more testable than the other, both have zero probability.) If increasing probability were our aim, we should prefer some weak consequence of a highly testable theory to the theory itself, since there is bound to be a consequence that is more probable than the theory. Thus the security-pole and the depth-pole of the Bacon-Descartes ideal pull in opposite directions, one calling for less, and the other for more, testable content.

If one compares what was happening *in* science from, say, Planck's discovery of the quantization of energy around 1900 to the discovery of nuclear fission in the 1930s, with what such philosophers as Mach, Duhem, Bridgman, and the members of the Vienna Circle were saying *about* science, one is struck by a remarkable contrast. It is as if Planck, Einstein, Rutherford, Bohr and others were inspired by the depth-pole of our ideal to penetrate to ever deeper layers of physical reality, whereas those philosophers of science were demanding that scientists should eschew depth and remain on the surface, at the phenomenal level. No doubt the various philosophers who waged this « antidepth war », as Mario Bunge called it¹⁸, each had his own set of reasons which differed somewhat from one to another. But my own diagnosis is that they intuitively perceived the antagonism between depth and security, and renounced the former in the hope of attaining the latter. This is particularly clear in the case of Moritz Schlick. In his *Allgemeine Erkenntnislehre* (1918) he had not yet perceived the antagonism between (A) and (B) and he strongly endorsed them both. As to (A) : he spoke of the « requirements for rigor and certainty » in science, of the « goal of abso-

17. Cf. K. R. POPPER, *op. cit. supra* n. 9, Section 83.

18. « The Maturation of Science », in Imre LAKATOS, Alan MUSGRAVE, eds, *Problems in the Philosophy of Science*, Amsterdam, North-Holland, 1968, p. 136.

lute certainty and precision » ; but he also conceded, as we saw earlier, that the truth of scientific hypotheses « is not absolutely guaranteed », adding that we must be content if their probability « assumes an extremely high value ». As to (B) : this book is characterized by a strong scientific realism : science penetrates beneath the surface to underlying realities. It tells us « about the interior of the sun, about electrons, about magnetic field strengths ». Indeed, science gets to the very essences of things : « Maxwell's equations disclose to us the "essence" of electricity, Einstein's equations the essence of gravitation. » But by 1931 Schlick had become acutely aware of the conflict between the goal of high probability and the goal of explanatory depth ; and he was anxious to solve Hume's problem. He continued to adhere resolutely to (A) and now totally repudiated (B) : the « laws » of science were now understood, not as genuine assertions about the world but as rules that license inferences from singular observation statements to other singular observation statements. (For references see my *Science and Scepticism*, section 4.35.)

Now if both the (A)-component and the (B)-components of the Bacon-Descartes ideal constituted possible aims for science which, though mutually exclusive, are separately feasible, then there would be no such thing as the optimum aim for science. We would face a choice between two aims neither of which dominates the other. But (A) is not a feasible aim for science. I mentioned earlier what I consider to be one insuperable difficulty confronting probabilism : a probabilist, whether or not he is a Bayesian, must judge the posterior probability of a hypothesis in the light of the total evidence ; but this is something he can never do. In *Science and Scepticism* I also go into what I call the « next-instance » thesis, according to which the predictive content of a theory should be taken to be, not the totality of its predictive implications, but only what it predicts for the next instance (or the next application, or the next year, say). Now if « next » is here used as in « The next number in the sequence 1,2,4,8,16... is always twice the previous number », it turns into a variable that extends over all future instances ; the theory's predictive content has not been cut back, and the previous sceptical arguments continue to apply. Then suppose that « next » is used as in « The next item on the agenda is the treasurer's report », so that it becomes a singular denoting term. Understood in this sense, the « next-instance » thesis would indeed effect a drastic reduction in the predictive content of a theory, making it like a bee that can only sting once. Would this drastic measure at least give the probabilist the desired result that evidence can raise the probability of the little that remains of the theory ? Carnap had claimed that instead of taking the roundabout way from evidence via the theory to a prediction yielded by it, we should go straight from the evi-

dence to the prediction ¹⁹. Now a good scientific theory makes *novel* predictions ; and there must be a time when the next instance is also the *first* instance of such a prediction. An example used by Hilary Putnam, in criticism of Carnap, was the prediction of the first nuclear explosion ²⁰. In his reply, Carnap conceded that the totality of the evidence available beforehand would not assign « a considerable value » to the probability of the prediction ²¹. But the question is not whether it would raise it to a considerable value, but whether it would raise it at all. And it is clear that it would not. It would be as if the evidence concerned red robins, black crows, white swans, etc., and the prediction were of a remarkable bird of a kind never hitherto observed.

My conclusion in this part of the book is that those antidepth philosophers of science who attempted to propitiate this almost universally accepted aim, gave an increasingly impoverished account of science in which one valuable feature after another (realism, universality) of science was given up, in the hope of rendering it probabilifiable. *And all these sacrifices were in vain* : the emasculated residue was still incapable of attaining the kind of probability that was sought.

If that is so, the question arises whether there is, within component (A), a feasible core that we can retrieve and combine with the (B)-components, thus creating a stronger aim that would dominate any other aim that is coherent and feasible, and hence would be the optimum aim for science. In the original, absolute version of the Bacon-Descartes ideal, component (A) called for scientific theories that are *certainly true*. Spelt out more fully, it said that a necessary and sufficient condition for it to be rational for person *X* to accept a scientific theory at time *t* is that the theory is certainly true for *X* at *t*. In its « progressivised » version, it called for scientific theories that are ever more *probably true*. Now if a scientific theory, however much de-ontologised and emasculated by positivist reinterpretation, cannot even have its probability *raised* by seemingly favourable evidence, then the next weakening of component (A) that is needed to render it feasible must presumably be to make it call for scientific theories that are possibly true. Spelt out more fully it would now say that a necessary, but not a sufficient, condition for it to be rational for a person *X* to accept a scientific theory at time *t* is that the theory is possibly true for *X* at *t*. But just what would this mean ? We might try saying that it means that *X* knows that the theory is internally consis-

19. Cf. R. CARNAP, *op. cit. supra* n. 5, p. 575-5.

20. « "Degree of Confirmation" and Inductive Logic », in Paul Arthur SCHILPP, ed., *The Philosophy of Rudolf Carnap*, La Salle, Open Court, 1963, p. 779.

21. *The Philosophy of Rudolf Carnap, op. cit. supra* n. 20, p. 988.

tent and that there are no conflicts between it and the evidence in his possession at t . But that would again render it an infeasible demand. It is impossible for X to be aware of all the logical implications of the theory, because there are infinitely many of these. And it is impossible for X to say of any large body of evidence E that E is the *total* evidence in his possession at t , for reasons given earlier. Thus it will always be possible that there is a consequence of this theory that does in fact conflict with the evidence in X 's possession, though X has not noticed this. Then should we say that a theory is possibly true for X at t if X does not know of any internal inconsistencies in it or of any conflicts between it and the evidence in his possession at t ? But that would render it too lax. Perhaps X is a lazy fellow who prefers not to search for internal inconsistencies or adverse evidence. If our modified version of (A) is to be neither too strong nor too weak, then it should say that it is a necessary condition for the rational acceptance of a scientific theory by X at t that X has not succeeded in discovering any internal inconsistencies in it, or evidence adverse to it, despite his best endeavours to do so.

Combining this modified version of (A) with the previous (B)-components results in the normative claim that, of all the competing scientific theories in a certain field that are possibly true in the above sense, the best is the one that is deepest and most unified, predictively powerful and exact. The question now is whether, if a certain theory is the best *corroborated* one in its set, it is thereby the *best* theory in the above sense.

7. — CORROBORATION AND RATIONAL PREFERENCE

My account of corroboration is derived from Popper's, though it differs over some details. In his theory, so-called « background knowledge » plays an essential role, since the measure of the severity of a test is given by $p(e, hb) - p(e, b)$, where e is a predicted test-result, h is the hypothesis, and b is background knowledge. Now if $p(e, b)$ is to have a determinate value we need to know exactly what b contains. But if we cannot know what our total evidence is, still less can we know what our total « background knowledge » is. I dispense with this concept and replace it with the idea of a historical record of tests in the field of the theory in question. A test on a predictive consequence of a theory is said to be *hard* if the consequence is empirically novel, in the sense that no previous test would have constituted a test of it if it had been formulated then. A test is *soft* if it can be regarded as a mere repetition of a previous test which this predictive consequence would have passed. And a test is *medium* if it is more stringent or searching than the previous tests that

this predictive consequence has passed (or would have passed if it had been formulated then).

I take the total testable content of a theory T , which I denote by $CT(T)$, to be the totality of all its singular predictive implications, that is, consequences of it that say that if such-and-such experimental conditions were realized, then such-and-such an observable effect would result. In my book I provide rules for what I call a « natural » axiomatization of a theory. And I call an axiom all of whose predicates are theoretical a fundamental assumption, and one, some or all of whose predicates are observational, an auxiliary assumption, of the theory. Now part of the total testable content of T may be entailed just by its auxiliary assumptions alone ; and I say that T can gain a corroboration from a favourable test-result only if some or all of its fundamental assumptions were needed for the derivation of the prediction that was tested. If we denote its auxiliary assumptions by A we may represent what I call the corroborable content of T by $CT(T) - CT(A)$. This expression designates all those singular predictive implications of T that are not entailed by its auxiliary assumptions alone.

I say that theory T_1 is better corroborated than theory T_2 if T_1 is unrefuted and no test-result has been more favourable to T_2 than to T_1 while at least one test-result has been more favourable to T_1 . A test-result is more favourable to T_1 if it corroborates T_1 without corroborating T_2 or if it refutes T_2 without refuting T_1 .

Suppose that T_1 is the best corroborated theory in its field. If it were the sole unrefuted survivor then it would obviously be the best or most preferable theory in its field, given our demand that, to be acceptable, a theory must be possibly true. So suppose that there is another unrefuted theory, say T_2 , in its field and that T_1 is better corroborated than T_2 . Can we conclude that T_1 is the best theory in its field at the present time, in the sense that it fulfils the optimum aim for science better than all its rivals ?

To answer this we need to know something about what tests have been made. After all, it *might* be that T_1 is better corroborated than T_2 because *no* tests have been made on T_2 while just one test, with a favourable outcome, has been made on T_1 . So I will introduce the following assumption. If either theory has excess corroborable content over the other, in the sense that it yields predictions in an area where the other theory is silent, then at least one test has been made on this excess content.

Since T_1 is better corroborated than T_2 there is at least one test-result that is more favourable to T_1 . This could mean either (i) that the result corroborated T_1 but not T_2 or (ii) that it refuted T_2 but not T_1 . However (ii) is here excluded because T_2 is unrefuted. So T_1 must have excess

corroborable content over T_2 . On the other hand, T_2 cannot have excess corroborable content over T_1 . For suppose that it did. Then by the above assumption there would have been at least one test on this excess content ; and this test would have resulted in either (i) a corroboration for T_2 but not for T_1 , or (ii) a refutation for T_2 . But (i) is ruled out by our supposition that T_1 is better corroborated and hence that no test-result has been less favourable to it than to T_2 , and (ii) is ruled out by our supposition that T_2 is unrefuted.

We may represent the corroborable content of these two theories by, respectively, $CT(T_1) - CT(A_1)$ and $CT(T_2) - CT(A_2)$. And as we have just seen, the former is larger than the latter. Now the main idea behind the technical elucidations in my book of the notion of one theory being deeper-and-more-unified than a rival is this : the deeper theory has the greater testable content, *not* because it is fitted out with a more powerful set of auxiliary assumptions, but because its *fundamental* assumptions (when married to appropriate auxiliary assumptions) are more fertile in predictive implications than are those of its rival. For this to be the case, it is not sufficient that the *total* testable content of T_1 is greater than that of T_2 (since that might have been achieved merely by more powerful auxiliary assumptions) ; but it is sufficient that the *corroborable* content of T_1 is greater than that of T_2 ; which is what we have here. So in the present case we can say that, while both T_1 and T_2 , being unrefuted and possibly true, satisfy our modified version of component (A) of the optimum aim for science, the better corroborated T_1 satisfies component (B1-2) better than T_2 . And if it does this it will automatically satisfy (B3-4) better : a deeper theory will also be wider, or have more explanatory and predictive power at the empirical level, although the converse does not hold (a theory may be wider but not deeper than a rival, for instance in virtue of more powerful auxiliary assumptions).

Thus on our assumption that at least one test has been made on any excess corroborable content that either of the two theories may have, we can conclude that the best corroborated theory is, at the present time, the best theory, that is, the theory that best fulfils the optimum aim for science.

Many philosophers assume that if Humean scepticism cannot be repelled, then there can be no scientific rationality ; for surely the aim of science is to establish theories, if not as true, then at least as more or less probably true ? This assumption plays into the hands of enemies of science, just because scientific theories cannot possibly be established as probably true. There is no valid rebuttal of Humean scepticism. Now it is not rational to set oneself an aim that simply cannot be fulfilled ; an aim must be feasible. So the first step, in rescuing the rationality of

science, must be to discard this infeasible element from a proposed aim for science. And once we discard it, we can reintroduce all those highly desirable elements that had been driven out, at its behest, during the anti-depth war. We can have a richly ambitious aim for science. And with the replacement of that old, impoverishing aim by this one, we restore the idea of the rationality of science. For this aim, though lofty, is sharply defined, and in such a way that we can normally tell which of several competing theories best fulfils it. Thus we can make rational preferences between existing theories, and we can know in advance what a new theory would need to achieve for it to be an advance on the best of the existing ones. The rationality of science is vindicated.

John WATKINS,
The London School of Economics.