THE LOGIC OF SCIENTIFIC DISCOVERY

KARL R. POPPER

BASIC BOOKS, INC.
NEW YORK

other empirical statements, to search for the forbidden. For he can test empirical statements only by trying to falsify them.

From an historical point of view, the emergence of indeterminist metaphysics is understandable enough. For a long time, physicists believed in determinist metaphysics. And because the logical situation was not fully understood, the failure of the various attempts to deduce the light spectra—which are statistical effects—from a mechanical model of the atom was bound to produce a crisis for determinism. Today we see clearly that this failure was inevitable, since it is impossible to deduce statistical laws from a non-statistical (mechanical) model of the atom. But at that time (about 1924, the time of the theory of Bohr, Kramers, and Slater) it could not but seem as if in the mechanism of each single atom, probabilities were taking the place of strict laws. The determinist edifice was wrecked—mainly because probability statements were expressed as formally singular statements. On the ruins of determinism, indeterminism rose, supported by Heisenberg's uncertainty principle. But it sprang, as we now see, from that same misunderstanding of the meaning of formally-singular probability statements.

The lesson of all this is that we should try to find strict laws—prohibitions—that can founder upon experience. Yet we should abstain from issuing prohibitions that draw limits to the possibilities of research.

CHAPTER X

CORROBORATION, OR HOW A THEORY STANDS UP TO TESTS

THEORIES are not verifiable, but they can be 'corroborated'.

The attempt has often been made to describe theories as being neither *true* nor *false*, but instead more or less *probable*. Inductive logic, more especially, has been developed as a logic which may ascribe not only the two values 'true' and 'false' to statements, but also degrees of probability; a type of logic which will here be called '*probability logic*'. According to those who believe in probability logic, induction should determine the degree of probability of a statement. And a principle of induction should either *make it sure* that the induced statement is 'probably valid' or else it should *make it probable*, in its turn—for the principle of induction might itself be only 'probably valid'. Yet in my view, the whole problem of the probability of hypotheses is misconceived. Instead of discussing the 'probability' of a hypothesis we should try to assess what tests, what trials, it has withstood; that is, we should try to assess how far it has been able to prove its fitness to survive by standing up to tests. In brief, we should try to assess how far it has been 'corroborated'.*1

*1 I introduced the terms '*corroboration*' ('*Bewährung*') and especially '*degree of corroboration*' ('*Grad der Bewährung*', '*Bewährungsgrad*') in my book because I wanted a *neutral* term to describe the degree to which a hypothesis has stood up to severe tests, and thus 'proved its mettle'. By 'neutral' I mean a term not prejudging the issue whether, by standing up to tests, the hypothesis becomes 'more probable', in the sense of the probability calculus. In other words, I introduced the term 'degree of corroboration' mainly in order to be able to discuss the problem whether or not 'degree of corroboration' could be indentified with 'probability' (either in a frequency sense or in the sense of Keynes, for example).

Carnap translated my term 'degree of corroboration' ('*Grad der Bewährung*'), which I had first introduced into the discussions of the Vienna Circle, as 'degree of confirmation'. (See his 'Testability and Meaning', in *Philosophy of Science* **3**, 1936; especially p. 427); and so the term 'degree of confirmation' soon became widely accepted. I did not like this term, because of some of its associations ('make firm'; 'establish firmly';

79. *Concerning the So-Called Verification of Hypotheses.*

The fact that theories are not verifiable has often been overlooked. People often say of a theory that it is verified when some of the predictions derived from it have been verified. They may perhaps admit that the verification is not completely impeccable from a logical point of view, or that a statement can never be finally established by establishing some of its consequences. But they are apt to look upon such objections as due to somewhat unnecessary scruples. It is quite true, they say, and even trivial, that we cannot know for certain whether the sun will rise tomorrow; but this uncertainty may be neglected: the fact that theories may not only be improved but that they can also be *falsified by new experiments* presents to the scientist a serious possibility which may at any moment become actual; but never yet has a theory had to be regarded as falsified owing to the sudden breakdown of a well-confirmed law. It never happens that old experiments one day yield new results. What happens is only that new experiments decide against an old theory. The old theory, even when it is superseded, often retains its validity as a kind of limiting case of the new theory; it still applies, at least with a high degree of approximation, in those cases in which it was successful before. In short, regularities which are directly testable by experiment do not change. Admittedly it is conceivable, or logically possible, that they might change; but this possibility is disregarded by empirical science and does not affect its methods. On the contrary, scientific method presupposes *the immutability of natural processes*, or the 'principle of the uniformity of nature'.

There is something to be said for the above argument, but it does not affect my thesis. It expresses the metaphysical faith in the existence of regularities in our world (a faith which I share, and without which practical action is hardly conceivable).[*1] Yet the question before us—

[*1] *Cf.* appendix *x, and also section *15 of my *Postscript.*

'put beyond doubt'; 'prove'; 'verify': 'to confirm' corresponds more closely to '*erhärten*' or '*bestätigen*' than to '*bewähren*'). I therefore proposed in a letter to Carnap (written, I think, about 1939) to use the term 'corroboration'. (This term had been suggested to me by Professor H. N. Parton.) But as Carnap declined my proposal, I fell in with his usage, thinking that words do not matter. This is why I myself used the term 'confirmation' for a time in a number of my publications.

Yet it turned out that I was mistaken: the associations of the word 'confirmation' did matter, unfortunately, and made themselves felt: 'degree of confirmation' was soon used—by Carnap himself—as a synonym (or 'explicans') of 'probability'. I have therefore now abandoned it in favour of 'degree of corroboration'. See also appendix *ix, and section *29 of my *Postscript.*

the question which makes the non-verifiability of theories significant in the present context—is on an altogether different plane. Consistently with my attitude towards other metaphysical questions, I abstain from arguing for or against faith in the existence of regularities in our world. But I shall try to show that *the non-verifiability of theories is methodologically important*. It is on this plane that I oppose the argument just advanced.

I shall therefore take up as relevant only one of the points of this argument—the reference to the so-called 'principle of the uniformity of nature'. This principle, it seems to me, expresses in a very superficial way an important methodological rule, and one which might be derived, with advantage, precisely from a consideration of the non-verifiability of theories.[*2]

Let us suppose that the sun will not rise tomorrow (and that we shall nevertheless continue to live, and also to pursue our scientific interests). Should such a thing occur, science would have to try to *explain* it, *i.e.* to derive it from laws. Existing theories would presumably require to be drastically revised. But the revised theories would not merely have to account for the new state of affairs: *our older experiences would also have to be derivable from them.* From the methodological point of view one sees that the principle of the uniformity of nature is here replaced by the postulate of *the invariance of natural laws*, with respect to both space and time. I think, therefore, that it would be a mistake to assert that natural regularities do not change. (This would be a kind of statement that can neither be argued against nor argued for.) What we should say is, rather, that it is part of our *definition* of natural laws if we postulate that they are to be invariant with respect to space and time; and also if we postulate that they are to have no exceptions. Thus from a methodological point of view, the possibility of falsifying a corroborated law is by no means without significance. It helps us to find out what we demand and expect from natural laws. And the 'principle of the uniformity of nature' can again be regarded as a metaphysical interpretation of a methodological rule—like its near relative, the 'law of causality'.

One attempt to replace metaphysical statements of this kind by principles of method leads to the 'principle of induction', supposed to govern the method of induction, and hence that of the verification

[*2] I mean the rule that any new system of hypotheses should yield, or explain, the old, corroborated, regularities. See also section *3 (third paragraph) of my *Postscript.*

of theories. But this attempt fails, for the principle of induction is itself metaphysical in character. As I have pointed out in section 1, the assumption that the principle of induction is empirical leads to an infinite regress. It could therefore only be introduced as a primitive proposition (or a postulate, or an axiom). This would perhaps not matter so much, were it not that the principle of induction would have in any case to be treated as a *non-falsifiable statement*. For if this principle—which is supposed to validate the inference of theories—were itself falsifiable, then it would be falsified with the first falsified theory, because this theory would then be a conclusion, derived with the help of the principle of induction; and this principle, as a premise, will of course be falsified by the *modus tollens* whenever a theory is falsified which was derived from it.*[3] But this means that a falsifiable principle of induction would be falsified anew with every advance made by science. It would be necessary, therefore, to introduce a principle of induction assumed not to be falsifiable. But this would amount to the misconceived notion of a synthetic statement which is *a priori* valid, *i.e.* an irrefutable statement about reality.

Thus if we try to turn our metaphysical faith in the uniformity of nature and in the verifiability of theories into a theory of knowledge based on inductive logic, we are left only with the choice between an infinite regress and *apriorism*.

80. *The Probability of a Hypothesis and the Probability of Events: Criticism of Probability Logic.*

Even if it is admitted that theories are never finally verified, may we not succeed in making them secure to a greater or lesser extent—more probable, or less so? After all, it might be possible that the question of *the probability of a hypothesis* could be reduced, say, to that of *the probability of events*, and thus be made susceptible to mathematical and logical handling.*[1]

Like inductive logic in general, the theory of the probability of

*[3] The premises of the derivation of the theory would (according to the inductivist view here discussed) consist of the principle of induction *and* of observation statements. But the latter are here tacitly assumed to be unshaken and reproducible, so that they cannot be made responsible for the failure of the theory.

*[1] The present section (80) contains mainly a criticism of the attempt (Reichenbach's) to interpret *the probability of hypotheses* in terms of *a frequency theory of the probability of events*. A criticism of Keynes's approach is contained in section 83.

hypotheses seems to have arisen through a confusion of psychological with logical questions. Admittedly, our subjective feelings of conviction are of different intensities, and the degree of confidence with which we await the fulfilment of a prediction and the further corroboration of a hypothesis is likely to depend, among other things, upon the way in which this hypothesis has stood up to tests so far—upon its past corroboration. But that these psychological questions do not belong to epistemology or methodology is pretty well acknowledged even by the believers in probability logic.*[2] They argue, however, that it is possible, on the basis of inductivist decisions, to ascribe degrees of probability to *the hypotheses themselves*; and further, that it is possible to reduce this concept to that of the probability of events.

The probability of a hypothesis is mostly regarded as merely a special case of the general problem of the *probability of a statement*; and this in turn is regarded as nothing but the problem of the *probability of an event*, expressed in a particular terminology. Thus we read in Reichenbach, for example: 'Whether we ascribe probability to statements or to events is only a matter of terminology. So far we have regarded it as a case of the probability of events that the probability of 1/6 has been assigned to the turning up of a certain face of a die. But we might just as well say that it is the *statement* "the face showing the 1 will turn up" which has been assigned the probability of 1/6.'[1]

This identification of the probability of events with the probability of statements may be better understood if we recall what was said in section 23. There the concept 'event' was defined as a class of singular statements. It must therefore also be permissible to speak of the *probability of statements* in place of the probability of events. So we can regard this as being merely a change of terminology: the reference-sequences are interpreted as sequences of statements. If we think of an 'alternative', or rather of its elements, as represented by statements, then we can describe the turning up of heads by the statement '*k* is heads', and its failure to turn up by the negation of this statement. In this way we obtain a sequence of statements of the form p_j, p_k, $\bar{p}_l$, p_m, $\bar{p}_n$, . . ., in which a statement p_i is sometimes characterized as 'true', and sometimes (by placing a bar over its name) as 'false'. Probability within an alternative can thus be interpreted as *the relative*

*[2] I am alluding here to the school of Reichenbach rather than to Keynes.

[1] Reichenbach, *Erkenntnis* **1**, 1930, p. 171 *f*.

'*truth-frequency*'[2] *of statements within a sequence of statements* (rather than as the relative frequency of a property).

If we like, we can call the concept of probability, so transformed, the 'probability of statements' or the 'probability of propositions'. And we can show a very close connection between this concept and the concept of 'truth'. For if the sequence of statements becomes shorter and shorter and in the end contains only one element, *i.e.* only *one single* statement, then the probability, or truth-frequency, of the sequence can assume only one of the two values 1 and 0, according to whether the single statement is true or false. The truth or falsity of a statement can thus be looked upon as a limiting case of probability; and conversely, probability can be regarded as a generalization of the concept of truth, in so far as it includes the latter as a limiting case. Finally, it is possible to define operations with truth-frequencies in such a way that the usual truth-operations of classical logic become limiting cases of these operations. And the calculus of these operations can be called '*probability logic*'.[3]

But can we really identify *the probability of hypotheses* with the probability of statements, defined in this manner, and thus indirectly with the probability of events? I believe that this identification is the result of a confusion. The idea is that the probability of a hypothesis, since it is obviously a kind of probability of a statement, must come under the head of 'probability of statements' *in the sense just defined*. But this conclusion turns out to be unwarranted; and the terminology is thus highly unsuitable. Perhaps after all it would be better never to use the expression 'probability of statements' if we have the probability of events in mind.[*2]

[2] According to Keynes, *A Treatise on Probability* (1921), p. 101 *ff.*, the expression 'truth-frequency' is due to Whitehead; *cf.* the next note.

[3] I am giving here an outline of the construction of the probability logic developed by Reichenbach (*Wahrscheinlichkeitslogik, Sitzungsberichte der Preussischen Akademie der Wissenschaften*, Physik.-mathem. Klasse **29**, 1932, p. 476 *ff.*) who follows E. L. Post (*American Journal of Mathematics* **43**, 1921, p. 184), and, at the same time, the frequency theory of von Mises. Whitehead's form of the frequency theory, discussed by Keynes, *op. cit.* p. 101 *ff.* is similar.

[*2] I still think (a) that the so-called 'probability of hypotheses' cannot be interpreted by a truth-frequency; (b) that it is better to call a probability defined by a relative frequency—whether a truth-frequency or the frequency of an event—the 'probability of an event'; (c) that the so-called 'probability of a hypothesis' (in the sense of its acceptability) is *not* a special case of the 'probability of statements'. And I should now regard the 'probability of statements' as one interpretation (the logical interpretation) among several possible interpretations of the formal calculus of probability, rather than as a truth-frequency. (*Cf.* appendices *ii, *iv, and *ix, and my *Postscript.*)

However this may be, I assert that the issues arising from the concept of a *probability of hypotheses* are not even touched by considerations based on probability logic. I assert that if one says of a hypothesis that it is not true but 'probable', then this statement can under *no* circumstances be translated into a statement about the probability of events.

For if one attempts to reduce the idea of a probability of hypotheses to that of a truth-frequency which uses the concept of a sequence of statements, then one is at once confronted with the question: *with reference to what sequence* of statements can a probability value be assigned to a hypothesis? Reichenbach identifies an 'assertion of natural science'—by which he means a scientific hypothesis—itself with a reference-sequence of statements. He says, '. . . the assertions of natural science, which are never singular statements, are in fact sequences of statements to which, strictly speaking, we must assign not the degree of probability 1 but a smaller probability value. It is therefore only probability logic which provides the logical form capable of strictly representing the concept of knowledge proper to natural science.'[4] Let us now try to follow up the suggestion that the hypotheses themselves are sequences of statements. One way of interpreting it would be to take, as the elements of such a sequence, the various singular statements which can contradict, or agree with, the hypothesis. The probability of this hypothesis would then be determined by the truth-frequency of those among these statements which agree with it. But this would give the hypothesis a probability of $\frac{1}{2}$ if, on the average, it is refuted by every second singular statement of this sequence! In order to escape from this devastating conclusion, we might try two more expedients.[*3] One would be to ascribe to the hypothesis a certain probability—perhaps not a very precise one—on the basis of an estimate of the ratio of all the tests passed by it to all the tests which have not yet been attempted. But this way too leads nowhere. For this estimate can, as it happens, be computed with precision, and the result is always that the probability is zero. And finally, we could try to base our estimate upon the ratio of those tests which led to a favourable result to those which

[4] Reichenbach, *Wahrscheinlichkeitslogik* (*op. cit.* p. 488), p. 15 of the reprint.

[*3] It is here assumed that we have by now made up our minds that whenever there is a clear-cut falsification, we will attribute to the hypothesis the probability zero, so that the discussion is now confined to those cases in which no clear-cut falsification has been obtained.

led to an indifferent result—*i.e.* one which did not produce a clear decision. (In this way one might indeed obtain something resembling a measure of the subjective feeling of confidence with which the experimenter views his results.) But this last expedient will not do either, even if we disregard the fact that with this kind of estimate we have strayed a long way from the concept of a truth-frequency, and that of a probability of events. (These concepts are based upon the ratio of the true statements to those which are false, and we must not, of course, equate an indifferent statement with one that is objectively false.) The reason why this last attempt fails too is that the suggested definition would make the probability of a hypothesis hopelessly subjective: the probability of a hypothesis would depend upon the training and skill of the experimenter rather than upon objectively reproducible and testable results.

But I think it is altogether impossible to accept the suggestion that a hypothesis can be taken to be a sequence of statements. It would be possible if universal statements had the form: 'For every value of *k* it is true that at the place *k* so-and-so occurs.' If universal statements had this form, then we could regard basic statements (those that contradict, or agree with, the universal statement) as elements of a sequence of statements—the sequence to be taken for the universal statement. But as we have seen (*cf.* sections 15 and 28), universal statements do not have this form. Basic statements are never derivable from universal statements alone.[*4] The latter cannot therefore be regarded as sequences of basic statements. If, however, we try to take into consideration the sequence of those negations of basic statements which *are* derivable from universal statements, then the estimate for *every* self-consistent hypothesis will lead to the same probability, namely 1. For we should then have to consider the ratio of the *non-falsified* negated basic statements which can be derived (or other derivable statements) to the *falsified* ones. This means that instead of

[*4] As explained in section 28 above, the singular statements which *can* be deduced from a theory—the 'instantial statements'—are not of the character of basic statements or of observation statements. If we nevertheless decide to take the sequence of these statements and base our probability upon the truth frequency within this sequence, then the probability will be always equal to 1, however often the theory may be falsified; for as has been shown in section 28, note *1, almost any theory is 'verified' by almost all instances (*i.e.* by almost all places *k*.) The discussion following here in the text contains a very similar argument—also based upon 'instantial statements' (*i.e.* negated basic statements)—designed to show that the probability of a hypothesis, if based upon these negated basic statements, would always be equal to one.

considering a truth frequency we should have to consider the complementary value of a falsity frequency. This value however would be equal to 1. For the class of derivable statements, and even the class of the derivable negations of basic statements, are both infinite; on the other hand, there cannot be more than at most a finite number of accepted falsifying basic statements. Thus even if we disregard the fact that universal statements are never sequences of statements, and even if we try to interpret them as something of the kind and to correlate with them sequences of completely decidable singular statements, even then we do not reach an acceptable result.

We have yet to examine another, quite different, possibility of explaining the probability of a hypothesis in terms of sequences of statements. It may be remembered that we have called a given singular occurrence 'probable' (in the sense of a 'formally singular probability statement') if it is an *element of a sequence* of occurrences with a certain probability. Similarly one might try to call a hypothesis 'probable' if it is an *element of a sequence of hypotheses* with a definite truth-frequency. But this attempt again fails—quite apart from the difficulty of determining the reference sequence (it can be chosen in many ways; *cf.* section 71). For we cannot speak of a truth-frequency within a sequence of hypotheses, simply because we can never know of a hypothesis whether it is true. If we *could* know this, then we should hardly need the concept of the probability of a hypothesis at all. Now we might try, as above, to take the complement of the falsity-frequency within a sequence of hypotheses as our starting point. But if, say, we define the probability of a hypothesis with the help of the ratio of the non-falsified to the falsified hypotheses of the sequence, then, as before, the probability of *every* hypothesis within *every infinite* reference sequence will be equal to 1. And even if a *finite* reference sequence were chosen we should be in no better position. For let us assume that we can ascribe to the elements of some (*finite*) sequence of hypotheses a degree of probability between 0 and 1 in accordance with this procedure—say, the value 3/4. (This can be done if we obtain the information that this or that hypothesis belonging to the sequence has been falsified.) In so far as these *falsified* hypotheses are elements of the sequence, we thus would have to ascribe to them, *just because of this information*, not the value 0, but 3/4. And in general, the probability of a hypothesis would decrease by $1/n$ in consequence of the information that it is false, where n is the number of hypotheses in the reference sequence. All this glaringly contradicts the programme of expressing, in terms of a '*probability of hypotheses*', the degree of reliability which we have to ascribe to a hypothesis in view of supporting or undermining evidence.

This seems to me to exhaust the possibilities of basing the concept of the probability of a hypothesis on that of the frequency of true statements (or the frequency of false ones), and thereby on the frequency theory of the probability of events.[*5]

I think we have to regard the attempt to identify the probability of a hypothesis with the probability of events as a complete failure. This conclusion is quite independent of whether we accept the claim (it is Reichenbach's) that *all hypotheses of physics* are 'in reality', or

[*5] One might summarize my foregoing attempts to make sense of Reichenbach's somewhat cryptic assertion that the probability of a hypothesis is to be measured by a truth frequency, as follows. (For a similar summary, with criticism, see the penultimate paragraph of appendix *i.)

Roughly, we can try two possible ways to define the probability of a theory. One is to count the number of experimentally testable statements belonging to the theory, and to determine the relative frequency of those which turn out to be true; this relative frequency can then be taken as a measure of the probability of a theory. We may call this a *probability of the first kind*. Secondly, we can consider the theory as an element of a class of ideological entities—say, of theories proposed by other scientists—and we can then determine the relative frequencies within this class. We may call this a *probability of the second kind*.

In my text I tried, further, to show that each of these two possibilities of making sense of Reichenbach's idea of truth frequency leads to results which must be quite unacceptable to adherents of the probability theory of induction.

Reichenbach replied to my criticism, not so much by defending his views as by attacking mine. In his paper on my book (*Erkenntnis* **5**, 1935, pp. 267–284), he said that 'the results of this book are completely untenable', and explained this by a failure of my 'method'—by my failure 'to think out all the consequences' of my conceptual system.

Section iv of his paper (pp. 274 *f.*) is devoted to our problem—the probability of hypotheses. It begins: 'In this connection, some remarks may be added about the probability of theories—remarks which should render more complete my so far all too brief communications of the subject, and which may remove a certain obscurity which still surrounds the issue.' After this follows a passage which forms the second paragraph of the present note, headed by the word 'Roughly' (the only word which I have added to Reichenbach's text).

Reichenbach remains silent about the fact that his attempt to remove 'the obscurity which still surrounds the issue' is but a summary—a rough one, admittedly—of some pages of the very book which he is attacking. Yet in spite of this silence I feel that I may take it as a great compliment from so experienced a writer on probability (who at the time of writing his reply to my book had two books and about a dozen papers on the subject to his credit) that he does accept the results of my endeavours to 'think out the consequences' of his 'all too brief communications on the subject'. This success of my endeavours was due, I believe, to a rule of 'method': that we should always try to clarify and to strengthen our opponent's position as much as possible before criticizing him, if we wish our criticism to be worth while.

'on closer examination' nothing but probability statements (about some average frequencies within sequences of observations which always show deviations from some mean value), or whether we are inclined to make a distinction between two different *types* of natural laws—between the 'deterministic' or 'precision' laws on the one hand, and the 'probability laws' or 'hypotheses of frequency' on the other. For both of these types are hypothetical assumptions which in their turn can never become 'probable': they can only be corroborated, in the sense that they can 'prove their mettle' under fire—the fire of our tests.

How are we to explain the fact that the believers in probability logic have reached an opposite view? Wherein lies the error made by Jeans when he writes—at first in a sense with which I can fully agree—that '. . . we can know nothing . . . *for certain*', but then goes on to say: 'At best we can only deal in *probabilities*. [And] the predictions of the new quantum theory agree so well [with the observations] that the odds in favour of the scheme having some correspondence with reality are *enormous*. Indeed, we may say the scheme is *almost certain* to be quantitatively true . . .'?[5]

Undoubtedly the commonest error consists in believing that hypothetical estimates of frequencies, that is to say, hypotheses regarding probabilities, can in their turn be only probable; or in other words, in ascribing to *hypotheses of probability* some degree of an alleged *probability of hypotheses*. We may be able to produce a persuasive argument in favour of this erroneous conclusion if we remember that hypotheses regarding probabilities are, as far as their logical form is concerned (and without reference to our methodological requirement of falsifiability), neither verifiable nor falsifiable. (*Cf.* sections 65 to 68.) They are not verifiable because they are universal statements, and they are not strictly falsifiable because they can never be logically contradicted by any basic statements. They are thus (as Reichenbach puts it) *completely undecidable*.[6] Now they can, as I have tried to show, *be better, or less well, 'confirmed'*, which is to say that they may agree more, or less, with accepted basic state-

[5] Jeans, *The New Background of Science* (1934), p. 58. (Only the words 'for certain' are italicized by Jeans.)

[6] Reichenbach, *Erkenntnis* **1**, 1930, p. 169 (*cf.* also Reichenbach's reply to my note in *Erkenntnis* **3**, 1933, p. 426 *f.*). Similar ideas about the degrees of probability or certainty of inductive knowledge occur very frequently (*cf.* for instance Russell, *Our Knowledge of the External World*, 1926, p. 225 *f.*, and *The Analysis of Matter*, 1927, pp. 141 and 398).

ments. This is the point where, it may appear, probability logic comes in. The symmetry between verifiability and falsifiability accepted by classical inductivist logic suggests the belief that it must be possible to correlate with these 'undecidable' probability statements some scale of degrees of validity, something like 'continuous degrees of probability whose unattainable upper and lower limits are truth and falsity',[7] to quote Reichenbach again. According to my view, however, probability statements, just because they are completely undecidable, are *metaphysical* unless we decide to make them falsifiable by accepting a methodological rule. Thus the simple result of their non-falsifiability is not that they can be better, or less well corroborated, but *that they cannot be empirically corroborated at all*. For otherwise—seeing that they rule out nothing, and are therefore compatible with every basic statement—they could be said to be 'corroborated' by *every arbitrarily chosen basic statement* (of any degree of composition) provided it describes the occurrence of some relevant instance.

I believe that physics uses probability statements only in the way which I have discussed at length in connection with the theory of probability; and more particularly that it uses probability assumptions, just like other hypotheses, as falsifiable statements. But I should decline to join in any dispute about how physicists 'in fact' proceed, since this must remain largely a matter of interpretation.

We have here quite a nice illustration of the contrast between my view and what I called, in section 10, the 'naturalistic' view. What can be shown is, first, the internal logical consistency of my view, and secondly, that it is free from those difficulties which beset other views. Admittedly it is impossible to prove that my view is correct, and a controversy with upholders of another logic of science may well be futile. All that can be shown is that my approach to this particular problem is a consequence of the conception of science for which I have been arguing.[*6]

81. *Inductive Logic and Probability Logic.*

The probability of hypotheses cannot be reduced to the probability

[7] Reichenbach, *Erkenntnis* **1**, 1930, p. 186 (*cf.* note 4 to section 1).

[*6] The last two paragraphs were provoked by the 'naturalistic' approach sometimes adopted by Reichenbach, Neurath, and others; *cf.* section 10, above.

of events. This is the conclusion which emerges from the examination carried out in the previous section. But might not a different approach lead to a satisfactory definition of the idea of a *probability of hypotheses*?

I do not believe that it is possible to construct a concept of the probability of hypotheses which may be interpreted as expressing a 'degree of validity' of the hypothesis, in analogy to the concepts 'true' and 'false' (and which, in addition, is sufficiently closely related to the concept 'objective probability', *i.e.* to relative frequency, to justify the use of the word 'probability').[1] Nevertheless, I will now, for the sake of argument, adopt the *supposition* that such a concept has in fact been successfully constructed, in order to raise the question: how would this affect the problem of induction?

Let us suppose that a certain hypothesis—say Schrödinger's theory—is recognized as 'probable' in some definite sense; either as 'probable to this or that numerical degree', or merely as 'probable', without specification of a degree. The statement that describes Schrödinger's theory as 'probable' we may call its *appraisal*.

An appraisal must, of course, be a synthetic statement—an assertion about 'reality'—in the same way as would be the statement 'Schrödinger's theory is true' or 'Schrödinger's theory is false'. All such statements obviously say something about the adequacy of the theory, and are thus certainly not tautological.[*1] They say that a theory

[1] (Added while the book was in proof.) It is conceivable that for estimating degrees of corroboration, one might find a formal system showing some limited formal analogies with the calculus of probability (*e.g.* with Bayes's theorem), without however haviing anything in common with the frequency theory. I am indebted to Dr. J. Hosiasson for suggesting this possibility to me. I am satisfied, however, that it is quite impossble to tackle *the problem of induction* by such methods with any hope of success. *See also note 3 to section *57 of my *Postscript*.

* Since 1938, I have upheld the view that 'to justify the use of the word probability', as my text puts it, we should have to show that the axioms of the formal calcuus are satisfied. (*Cf.* appendices *ii to *v, and especially section *28 of my *Postscript*.) This would of course include the satisfaction of Bayes's theorem. As to the formal analogies between Bayes's theorem on *probability* and certain theorems on *degree of corroboration*, see appendix *ix, point 9 (vii) of the first note, and points (12) and (13) of section *32 of my *Postscript*.

[*1] The probability statement '$p(S,e) = r$', in words, 'Schrödinger's theory, given the evidence e, has the probability r'—a statement of relative or conditional logical probability—may certainly be tautological (provided the values of e and r are chosen so as to fit each other: if e consists only of observational reports, r will have to equal zero in a sufficiently large universe). But the 'appraisal', in our sense, would have a different form (see section 84, below, especially the text to note *2)—for example, the following: $p_k(S) = r$, where k is today's date; or in words: 'Schrödinger's theory has *today* (in view of the actual total evidence now available) a probability of r.' In order to obtain this assessment, $p_k(S) = r$, from (i) the tautological statement of relative probability

is adequate or inadequate, or that it is adequate in some degree. Further, an appraisal of Schrödinger's theory must be a *non-verifiable* synthetic statement, just like the theory itself. For the 'probability' of a theory—that is, the probability that the theory will remain acceptable—cannot, it appears, be deduced from basic statements *with finality*. Therefore we are forced to ask: How can the appraisal be justified? How can it be tested? (Thus the problem of induction arises again; see section 1.)

As to the appraisal itself, this may either be asserted to be 'true', or it may, in its turn, be said to be 'probable'. If it is regarded as 'true' then it must be a *true synthetic statement* which has not been empirically verified—a synthetic statement which is *a priori* true. If it is regarded as 'probable', then we need a *new* appraisal: an appraisal of the appraisal, as it were, and therefore an appraisal on a higher level. But this means that we are caught up in an infinite regress. The appeal to the probability of the hypothesis is unable to improve the precarious logical situation of inductive logic.

Most of those who believe in probability logic uphold the view that the appraisal is arrived at by means of a 'principle of induction' which ascribes probabilities to the induced hypotheses. But if they ascribe a probability to this principle of induction in its turn, then the infinite regress continues. If on the other hand they ascribe 'truth' to it then they are left with the choice between infinite regress and *a priorism*. 'Once and for all', says Heymans, 'the theory of probability is incapable of explaining inductive arguments; for precisely the same problem which lurks in the one also lurks in the other (in the empirical application of probability theory). In both cases the

$p(S,e) = r$, and (ii) the statement 'e is the total evidence available today', we must apply a *principle of inference* (called the 'rule of absolution' in my *Postscript*, sections *43 and *51). This principle of inference looks very much like the *modus ponens*, and it may therefore seem that it should be taken as analytic. But if we take it to be analytic, then this amounts to the decision to consider p_k as *defined* by (i) and (ii), or at any rate as meaning *no more* than do (i) and (ii) together; but in this case, p_k cannot be interpreted as being of any practical significance: it *certainly* cannot be interpreted as a practical measure of acceptability. This is best seen if we consider that in a sufficiently large universe, $p_k(t,e) \approx 0$ for *every* universal theory t, provided e consists only of singular statements. (*Cf.* appendices, *vii and *viii.) But in practice, we certainly do accept some theories and reject others.

If, on the other hand, we interpret p_k as *degree of adequacy or acceptability*, then the principle of inference mentioned—the 'rule of absolution' (which, on this interpretation, becomes a typical example of a 'principle of induction')—is simply *false*, and therefore clearly non-analytic.

conclusion goes beyond what is given in the premises.'[2] Thus nothing is gained by replacing the word 'true' by the word 'probable', and the word 'false' by the word 'improbable'. Only if *the asymmetry between verification and falsification* is taken into account—that asymmetry which results from the logical relation between theories and basic statements—is it possible to avoid the pitfalls of the problem of induction.

Believers in probability logic may try to meet my criticism by asserting that it springs from a mentality which is 'tied to the framework of classical logic', and which is therefore incapable of following the methods of reasoning employed by probability logic. I freely admit that I am incapable of following these methods of reasoning.

82. *The Positive Theory of Corroboration: How a Hypothesis may 'Prove its Mettle'.*

Cannot the objections I have just been advancing against the probability theory of induction be turned, perhaps, against my own view? It might well seem that they can; for these objections are based on the idea of an *appraisal*. And clearly, I have to use this idea too. I speak of the '*corroboration*' of a theory; and corroboration can only be expressed as an appraisal. (In this respect there is no difference between corroboration and probability.) Moreover, I too hold that hypotheses cannot be asserted to be 'true' statements, but that they are 'provisional conjectures' (or something of the sort); and this view, too, can only be expressed by way of an appraisal of these hypotheses.

The second part of this objection can easily be answered. The appraisal of hypotheses which indeed I am compelled to make use of, and which describes them as 'provisional conjectures' (or something of the sort) has the status of a *tautology*. Thus it does not give rise

[2] Heymans, *Gesetze und Elemente des wissenschaftlichen Denkens* (1890, 1894), p. 290 f.; * third edition, 1915, p. 272. Heymans's argument was anticipated by Hume in his anonymous pamphlet, *An Abstract of a Book lately published entitled A Treatise of Human Nature*, 1740. I have little doubt that Heymans did not know this pamphlet which was re-discovered and attributed to Hume by J. M. Keynes and P. Sraffa, and published by them in 1938. I knew neither of Hume's nor of Heymans's anticipation of my arguments against the probabilistic theory of induction when I presented them in 1931 in an earlier book, still unpublished, which was read by several members of the Vienna Circle. The fact that Heymans's passage had been anticipated by Hume was pointed out to me by J. O. Wisdom; *cf.* his *Foundations of Inference in Natural Science*, 1952, p. 218. Hume's passage is quoted below, in appendix *vii, text to footnote 6.

to difficulties of the type to which inductive logic gives rise. For this description only paraphrases or interprets the assertion (to which it is equivalent by definition) that strictly universal statements, *i.e.* theories, cannot be derived from singular statements.

The position is similar as regards the first part of the objection which concerns appraisals stating that a theory is corroborated. The appraisal of the corroboration is not a hypothesis, but can be derived if we are given the theory as well as the accepted basic statements. It asserts the fact that these basic statements do not contradict the theory, and it does this with due regard to the degree of testability of the theory, and to the severity of the tests to which the theory has been subjected, up to a stated period of time.

We say that a theory is 'corroborated' so long as it stands up to these tests. The appraisal which asserts corroboration (the corroborative appraisal) establishes certain fundamental relations, *viz.* compatibility and incompatibility. We regard incompatibility as falsification of the theory. But compatibility alone must not make us attribute to the theory a positive degree of corroboration: the mere fact that a theory has not yet been falsified can obviously not be regarded as sufficient. For nothing is easier than to construct any number of theoretical systems which are compatible with any given system of accepted basic statements. (This remark applies also to all 'metaphysical' systems.)

It might perhaps be suggested that a theory should be accorded some positive degree of corroboration if it is compatible with the system of accepted basic statements, and if, in addition, part of this system can be derived from the theory. Or, considering that basic statements are not derivable from a purely theoretical system (though their negations may be so derivable), one might suggest that the following rule should be adopted: a theory is to be accorded a positive degree of corroboration if it is compatible with the accepted basic statements and if, in addition, a non-empty sub-class of these basic statements is derivable from the theory in conjunction with the other accepted basic statements.[*1]

[*1] The tentative definition of 'positively corroborated' here given (but rejected as insufficient in the next paragraph of the text because it does not explicitly refer to the results of severe tests, *i.e.* of attempted refutations) is of interest in at least two ways. First, it is closely related to my criterion of demarcation, especially to that formulation of it to which I have attached note *1 to section 21. In fact, the two agree except for the restriction to *accepted* basic statements which forms part of the present definition.

I have no serious objections to this last formulation, except that it seems to me insufficient for an adequate characterization of the positive degree of corroboration of a theory. For we wish to speak of theories as being better, or less well, corroborated. But the *degree of corroboration* of a theory can surely not be established simply by counting the number of the corroborating instances, *i.e.* the accepted basic statements which are derivable in the way indicated. For it may happen that one theory appears to be far less well corroborated than another one, even though we have derived very many basic statements with its help, and only a few with the help of the second. As an example we might compare the hypothesis 'All crows are black' with the hypothesis (mentioned in section 37) 'the electronic charge has the value determined by Millikan'. Although in the case of a hypothesis of the former kind, we have presumably encountered many more corroborative basic statements, we shall nevertheless judge Millikan's hypothesis to be the better corroborated of the two.

This shows that it is not so much the number of corroborating instances which determines the degree of corroboration as *the severity of the various tests* to which the hypothesis in question can be, and has been, subjected. But the severity of the tests, in its turn, depends upon the *degree of testability*, and thus upon the simplicity of the hypothesis: the hypothesis which is falsifiable in a higher degree, or the simpler hypothesis, is also the one which is corroborable in a higher degree.[1]

[1] This is another point in which there is agreement between my view of simplicity and Weyl's; *cf.* note 7 to section 42. *This agreement is a consequence of the view, due to Jeffreys, Wrinch, and Weyl (*cf.* note 7 to section 42), that the paucity of the parameters of a function can be used as a measure of its simplicity, taken in conjunction with my view (*cf.* sections 38 ff.) that the paucity of the parameters can be used as a measure of testability or improbability—a view rejected by these authors. (See also notes *1 and *2 to sections 43.)

Thus if we omit this restriction, the present definition turns into my criterion of demarcation.

Secondly, if instead of omitting this restriction we restrict the class of the *derived* accepted basic statements further, by demanding that they should be accepted as the results of sincere attempts to refute the theory, then our definition becomes an adequate definition of 'positively corroborated', though not, of course, of 'degree of corroboration'. The argument supporting this claim is implicit in the text here following. Moreover, the basic statements so accepted may be described as 'corroborating statements' of the theory.

It should be noted that 'instantial statements' (*i.e.* negated basic statements; see section 28) cannot be adequately described as corroborating or confirming statements of the theory which they instantiate, owing to the fact that we know that *every universal law is instantiated* almost everywhere, as indicated in note *1 to section 28. (See also note *4 to section 80, and text.)

Of course, the degree of corroboration actually attained does not depend *only* on the degree of falsifiability: a statement may be falsifiable to a high degree yet it may be only slightly corroborated, or it may in fact be falsified. And it may perhaps, without being falsified, be superseded by a better testable theory from which it—or a sufficiently close approximation to it—can be deduced. (In this case too its degree of corroboration is lowered.)

The degree of corroboration of two statements may not be comparable in all cases, any more than the degree of falsifiability: we cannot define a numerically calculable degree of corroboration, but can speak only roughly in terms of positive degrees of corroboration, negative degrees of corroboration, and so forth.[*2] Yet we can lay down various rules; for instance the rule that we shall not continue to accord a positive degree of corroboration to a theory which has been falsified by an inter-subjectively testable experiment based upon a falsifying hypothesis (*cf.* sections 8 and 22). (We may, however, under certain circumstances accord a positive degree of corroboration to another theory, even though it follows a kindred line of thought. An example is Einstein's photon theory, with its kinship to Newton's corpuscular theory of light.) In general we regard an inter-subjectively testable falsification as final (provided it is well tested): this is the way in which the asymmetry between verification and falsification of theories makes itself felt. Each of these methodological points contributes in its own peculiar way to the historical development of science as a process of step by step approximations. A corroborative appraisal made at a later date—that is, an appraisal made after new basic statements have been added to those already accepted—can replace a positive degree of corroboration by a negative one, but not *vice versa*. And although I believe that in the history of science it is always the theory and not the experiment, always the idea and not the observation, which opens up the way to new knowledge, I also believe that it is always the experiment which saves us from following a track that leads nowhere: which helps us out of the rut, and which challenges us to find a new way.

*2 As far as practical application to existing theories goes, this seems to me still correct; but I think now that it is possible to define 'degree of corroboration' in such a way that we can *compare* degrees of corroboration (for example, those of Newton's and of Einstein's theory of gravity). Moreover, this definition makes it even possible to attribute numerical degrees of corroboration to statistical hypotheses, and perhaps even to other statements *provided* we can attribute degrees of (absolute and relative) logical probability to them. See also appendix *ix.

Thus the degree of falsifiability or of simplicity of a theory enters into the appraisal of its corroboration. And this appraisal may be regarded as one of the logical relations between the theory and the accepted basic statements: as an appraisal that takes into consideration the severity of the tests to which the theory has been subjected.

83. *Corroborability, Testability, and Logical Probability.*[*1]

In appraising the degree of corroboration of a theory we take into account its degree of falsifiability. A theory can be the better corroborated the better testable it is. Testability, however, is converse to the concept of *logical probability*, so that we can also say that an appraisal of corroboration takes into account the logical probability of the statement in question. And this, in turn, as was shown in section 72, is related to the concept of objective probability—the probability of events. Thus by taking logical probability into account the concept of corroboration is linked, even if perhaps only indirectly and loosely, with that of the probability of events. The idea may occur to us that there is perhaps a connection here with the doctrine of the probability of hypotheses criticized above.

When trying to appraise the degree of corroboration of a theory we may reason somewhat as follows. Its degree of corroboration will increase with the number of its corroborating instances. Here we usually accord to the first corroborating instances far greater importance than to later ones: once a theory is well corroborated, further instances raise its degree of corroboration only very little. This rule however does not hold good if these new instances are very different from the earlier ones, that is if they corroborate the theory in a *new field of application*. In this case, they may increase the degree of corroboration very considerably. The degree of corroboration of a theory which has a higher degree of universality can thus be greater than that of a theory which has a lower degree of universality (and therefore a lower degree of falsifiability). In a similar way, theories of a higher degree of precision can be better corroborated than less precise ones. One of the reasons why we do not accord a positive degree of corroboration to the typical prophecies of palmists and soothsayers is that their pre-

*1 If the terminology is accepted which I first explained in my note in *Mind*, 1938, then the word 'absolute' should be inserted here throughout (as in sections 34, etc.) before 'logical probability' (in contradistinction to 'relative' or 'conditional' logical probability); *cf.* appendices *ii, *iv, and *ix.

dictions are so cautious and imprecise that the logical probability of their being correct is extremely high. And if we are told that more precise and thus logically less probable predictions of this kind have been successful, then it is not, as a rule, their success that we are inclined to doubt so much as their alleged logical improbability: since we tend to believe that such prophecies are non-corroborable, we also tend to argue in such cases from their low degree of corroborability to their low degree of testability.

If we compare these views of mine with what is implicit in (inductive) probability logic, we get a truly remarkable result. According to my view, the corroborability of a theory—and also the degree of corroboration of a theory which has in fact passed severe tests, stand both, as it were,[*2] in inverse ratio to its logical probability; for they both increase with its degree of testability and simplicity. *But the view implied by probability logic is the precise opposite of this.* Its upholders let the probability of a hypothesis increase in *direct proportion* to its logical probability—although there is no doubt that they *intend* their 'probability of a hypothesis' to stand for much the same thing that I try to indicate by 'degree of corroboration'.[*3]

[*2] I said in the text '*as it were*': I did so because I did not really believe in numerical (absolute) logical probabilities. In consequence of this, I wavered, when writing the text, between the view that the degree of corroborability is *complementary* to (absolute) logical probability and the view that it is inversely proportional; or in other words, between a definition of $C(g)$, *i.e.* the degree of corroborability, by $C(g) = 1 - P(g)$ which would make *corroborability equal to content*, and by $C(g) = 1/P(g)$, where $P(g)$ is the absolute logical probability of g. In fact, definitions may be adopted which lead to either of these consequences, and both ways seem fairly satisfactory on intuitive grounds; this explains, perhaps, my wavering. There are strong reasons in favour of the first method, or else of a logarithmic scale applied to the second method. See appendix *ix.

[*3] The last lines of this paragraph, especially from the italicized sentence on (it was not italicized in the original) contain the crucial point of my criticism of the probability theory of induction. The point may be summarized as follows.

We want *simple* hypotheses—hypotheses of a high *content*, a high degree of *testability*. These are also the highly *corroborable* hypotheses, for the degree of corroboration of a hypothesis depends mainly upon the severity of its tests, and thus upon its testability. Now we know that testability is the same as high (absolute) logical *improbability*, or low (absolute) logical *probability*.

But if two hypotheses, h_1 and h_2, are comparable with respect to their content, and thus with respect to their (absolute) logical probability, then the following holds: let the (absolute) logical probability of h_1 be smaller than that of h_2. Then, whatever the evidence e, the (relative) logical probability of h_1 given e can never exceed that of h_2 given e. Thus the *better testable and better corroborable hypothesis can never obtain a higher probability, on the given evidence, than the less testable one.* But this entails that *degree of corroboration cannot be the same as probability.*

This is the crucial result. My later remarks in the text merely draw the conclusion from it: if you value high probability, you must say very little—or better still, nothing at all: tautologies will always retain the highest probability.

Among those who argue in this way is Keynes who uses the expression '*a priori* probability' for what I call 'logical probability'. (See note 1 to section 34.) He makes the following perfectly accurate remark[1] regarding a 'generalization' g (*i.e.* a hypothesis) with the 'condition' or antecedent or protasis φ and the 'conclusion' or consequent or apodosis f: 'The more comprehensive the condition φ and the less comprehensive the conclusion f, the greater *à priori*[*4] probability do we attribute to the generalization g. With every increase in φ this probability increases, and with every increase in f it will diminish.' This, as I said, is perfectly accurate, even though Keynes does not draw a sharp distinction[*5] between what he calls the 'probability of a generalization'—corresponding to what is here called the 'probability of a hypothesis'—and its '*a priori* probability'. That Keynes nevertheless intends by his 'probability' the same as I do by my 'corroboration' may be seen from the fact that his 'probability' rises with the number of corroborating instances, and also (most important) with the increase of diversity among them. (But Keynes overlooks the fact that theories whose corroborating instances belong to widely different fields of application will usually have a correspondingly high degree of universality. Hence his two requirements for obtaining a high probability—the least possible universality and the greatest possible diversity of corroborating instances—will be, as a rule, incompatible.)

Expressed in my terminology, Keynes's theory implies that corroboration (or the probability of hypotheses) *decreases* with testability.

[1] Keynes, *A Treatise on Probability* (1921), pp. 224 *f.* Keynes's condition φ and conclusion f correspond (*cf.* note 6 to section 14) to our conditioning statement function φ and our consequence statement function f; *cf.* also section 36. It should be noticed that Keynes called the condition or the conclusion *more comprehensive* if its *content*, or its intension, rather than its extension, is the greater. (I am alluding to the inverse relationship holding between the intension and the extension of a term.)

[*4] Keynes persistently follows some other eminent Cambridge logicians in writing '*à priori*' and '*à posteriori*'; one can only say, *à propos de rien*—unless, perhaps, apropos of '*à propos*'.

[*5] Keynes does, in fact, allow for the distinction between the *a priori* (or 'absolute logical', as I now call it) probability of the 'generalization' g and its probability with respect to a given piece of evidence h, and to this extent, my statement in the text needs correction. (He makes the distinction by assuming, correctly though perhaps only implicitly—see p. 225 of the *Treatise*—that if $\varphi = \varphi_1\varphi_2$, and $f = f_1 f_2$, then the *a priori* probabilities of the various g are: $g(\varphi, f_1) \geqslant g(\varphi, f) \geqslant g(\varphi_1, f)$.) And he correctly *proves* that the (*a posteriori*) probabilities of these hypotheses g (with respect to any given piece of evidence h) are related in the same way as their *a priori* probabilities. Thus he proves that the probabilities are related like the (absolute) logical probabilities; while my cardinal point was, and still is, that their degree of corroborability and of corroboration are related in the opposite way.

He is led to this view by his belief in inductive logic.*[6] For it is the tendency of inductive logic to make scientific hypotheses as *certain* as possible. Scientific significance is assigned to the various hypotheses only to the extent to which they can be justified by experience. A theory is regarded as scientifically valuable only because of the close *logical proximity* (*cf.* note 2 to section 48 and text) between the theory and empirical statements. But this means nothing else than that the *content* of the theory must go *as little as possible* beyond what is empirically established.*[7] This view is closely connected with a tendency to deny the value of prediction. 'The peculiar virtue of prediction' Keynes writes[2] '. . . is altogether imaginary. The number of instances examined and the analogy between them are the essential points, and the question as to whether a particular hypothesis happens to be propounded before or after their examination is quite irrelevant.' In reference to hypotheses which have been '*a priori* proposed'—that is, proposed before we had sufficient support for them on inductive grounds—Keynes writes: '. . . if it is a mere guess, the lucky fact of its preceding some or all of the cases which verify it adds nothing whatever to its value.' This view of prediction is certainly consistent. But it makes one wonder why we should ever have to generalize at all. What possible reason can there be for constructing all these theories and hypotheses? The standpoint of inductive logic makes these activities quite incomprehensible. If what we value most is the securest knowledge available—and if predictions as such contribute nothing towards corroboration—why then may we not rest content with our basic statements?*[8]

Another view which gives rise to very similar questions is that

*[6] See my *Postscript*, chapter *ii. In my theory of corroboration—in direct opposition to Keynes's, Jeffreys's, and Carnap's theories of probability—corroboration does not *decrease* with testability, but tends to *increase* with it.

*[7] This may also be expressed by the unacceptable rule: 'Always choose the hypothesis which is most *ad hoc*!'

[2] Keynes, *op. cit.*, p. 305.

*[8] Carnap, in his *Logical Foundations of Probability* (1950) believes in the *practical* value of predictions; nevertheless, he draws part of the conclusion here mentioned—that we might be content with our basic statements. For he says that theories (he speaks of 'laws') are 'not indispensable' for science—not even for making predictions: we can manage throughout with singular statements. 'Nevertheless', he writes (p. 575) 'it is expedient, of course, to state universal laws in books on physics, biology, psychology, *etc.*' But the question is not one of expediency—it is one of scientific curiosity. *Some scientists want to explain the world*: their aim is to find satisfactory explanatory theories—well testable, *i.e.* simple theories—and to test them. (See also appendix *x and section *15 of my *Postscript*.)

of Kaila.[3] Whilst I believe that it is the simple theories, and those which make little use of auxiliary hypotheses (*cf.* section 46) which can be well corroborated, just because of their logical improbability, Kaila interprets the situation in precisely the opposite way, on grounds similar to Keynes's. He too sees that we usually ascribe a high probability (in our terminology, a high 'probability of hypotheses') to *simple* theories, and especially to those needing few auxiliary hypotheses. But his reasons are the opposite of mine. He does not, as I do, ascribe a high probability to such theories because they are severely testable, or logically improbable; that is to say because they have, *a priori* as it were, *many opportunities of clashing with basic statements*. On the contrary he ascribes this high probability to simple theories with few auxiliary hypotheses because he believes that a system consisting of *few* hypotheses will, *a priori*, have *fewer* opportunities of clashing with reality than a system consisting of many hypotheses. Here again one wonders why we should ever bother to construct these adventurous theories. If we shrink from conflict with reality, why invite it by making assertions? Our aim being security, our safest course would be to adopt a system *without* hypotheses.

My own rule which requires that auxiliary hypotheses shall be used as sparingly as possible (the 'principle of parsimony in the use of hypotheses') has nothing whatever in common with considerations such as Kaila's. I am not interested in merely keeping down the number of our statements: I am interested in their *simplicity in the sense of high testability*. It is this interest which leads, on the one hand, to my rule that auxiliary hypotheses should be used as sparingly as possible, and on the other hand, to my demand that the number of our axioms—of our most fundamental hypotheses—should be kept down. For this latter point arises out of the demand that statements of a high level of universality should be chosen, and that a system consisting of many 'axioms' should, if possible, be deduced from (and thus explained by) one with fewer 'axioms', and with axioms of a higher level of universality.

84. *Remarks Concerning the Use of the Concepts 'True' and 'Corroborated'.*

In the logic of science here outlined it is possible to avoid using

[3] Kaila, *Die Principien der Wahrscheinlichkeitslogik* (*Annales Universitatis Aboensis*, Turku 1926), p. 140.

the concepts 'true' and 'false'.*[1] Their place may be taken by logical considerations about derivability relations. Thus we need not say: 'The prediction p is true provided the theory t and the basic statement b are true.' We may say, instead, that the statement p follows from the (non-contradictory) conjunction of t and b. The falsification of a theory may be described in a similar way. We need not say that the theory is 'false', but we may say instead that it is contradicted by a certain set of accepted basic statements. Nor need we say of basic statements that they are 'true' or 'false', for we may interpret their acceptance as the result of a conventional decision, and the accepted statements as results of this decision.

This certainly does not mean that we are forbidden to use the concepts 'true' and 'false', or that their use creates any particular difficulty. The very fact that we can avoid them shows that they cannot give rise to any new fundamental problem. The use of the concepts 'true' and 'false' is quite analogous to the use of such concepts as

*[1] Not long after this was written, I had the good fortune to meet Alfred Tarski who explained to me the fundamental ideas of his theory of truth. It is a great pity that this theory—one of the two great discoveries in the field of logic made since *Principia Mathematica*—is still often misunderstood and misrepresented. It cannot be too strongly emphasized that Tarski's idea of truth (for whose definition with respect to formalized languages Tarski gave a method) is the same idea which Aristotle had in mind and indeed most people (except pragmatists): the idea that *truth is correspondence with the facts* (or with reality). But what can we possibly mean if we say of a *statement* that it corresponds with the *facts* (or with reality)? Once we realize that this correspondence cannot be one of structural similarity, the task of elucidating this correspondence seems hopeless; and as a consequence, we may become suspicious of the concept of truth, and prefer not to use it. Tarski solved this apparently hopeless problem (with respect to formalized languages), by reducing the unmanageable idea of correspondence to a simpler idea (that of 'satisfaction' or 'fulfilment').

Owing to Tarski's teaching, I am no longer hesitant in speaking of 'truth' and 'falsity'. And like everybody else's views (unless he is a pragmatist), my views turned out, as a matter of course, to be consistent with Tarski's theory of absolute truth. Thus although my views on formal logic and its philosophy were revolutionized by Tarski's theory, my views on science and its philosophy were unaffected.

The current criticism of Tarski's theory seems to me completely off the mark. It is said that his definition is artificial and complex; but since he defines truth with respect to formalized languages, it has to be based on the definition of a well-formed formula in such a language; and it is of precisely the same degree of 'artificiality' or 'complexity' as this definition. It is also said that only propositions or statements can be true or false, but not sentences. Perhaps 'sentence' was not a good translation of Tarski's original terminology. (I personally prefer to speak of 'statement' rather than of 'sentence'; see for example my 'Note on Tarski's Definition of Truth', *Mind* **64**, 1955, p. 388, footnote 1.) But Tarski himself made it perfectly clear that an uninterpreted formula (or a string of symbols) cannot be said to be true or false, and that these terms only apply to interpreted formulae—to '*meaningful* sentences' (as the translation has it). Improvements in terminology are always welcome; but it is sheer obscurantism to criticize a theory on terminological grounds.

'*tautology*', '*contradiction*', '*conjunction*', '*implication*' and others of the kind. These are non-empirical concepts, logical concepts.[1] They describe or appraise a statement irrespective of any changes in the empirical world. Whilst we assume that the properties of physical objects (of 'genidentical' objects in Lewin's sense) change with the passage of time, we decide to use these logical predicates in such a way that the logical properties of statements become timeless: if a statement is a tautology, then it is a tautology once and for all. This same timelessness we also attach to the concepts 'true' and 'false', in agreement with common usage. It is not common usage to say of a statement that it was perfectly true yesterday but has become false today. If yesterday we appraised a statement as true which today we appraise as false, then we implicitly assert today that *we were mistaken yesterday*; that the statement was false even yesterday—timelessly false—but that we erroneously 'took it for true'.

Here one can see very clearly the difference between truth and corroboration. The appraisal of a statement as corroborated or as not coroborated is also a logical appraisal and therefore also timeless; for it asserts that a certain logical relation holds between a theoretical system and some system of accepted basic statements. But we can never simply say of a statement that it is as such, or in itself, 'corroborated' (in the way in which we may say that it is 'true'). We can only say that it is *corroborated with respect to some system of basic statements*—a system accepted up to a particular point in time. 'The corroboration which a theory has received up to yesterday' is *logically not identical* with 'the corroboration which a theory has received up to today'. Thus we must attach a subscript, as it were, to every appraisal of corroboration—a subscript characterizing the system of basic statements to which the corroboration relates (for example, by the date of its acceptance).*[2]

Corroboration is therefore not a 'truth value'; that is, it cannot be placed on a par with the concepts 'true' and 'false' (which are free from temporal subscripts); for to one and the same statement there may be any number of different corroboration values, of which indeed all can be 'correct' or 'true' at the same time. For they are values which are logically derivable from the theory and the various sets of basic statements accepted at various times.

(Added in 1934 in proof.)

[1] Carnap would probably say 'syntactical concepts' (*cf.* his *Logical Syntax of Language*).

*[2] *Cf.* note *1 to section 81.

The above remarks may also help to elucidate the contrast between my views and those of the pragmatists who propose to *define 'truth' in terms of the success of a theory—and thus of its usefulness, or of its confirmation or of its corroboration.* If their intention is merely to assert that a logical appraisal of the success of a theory can be no more than an appraisal of its corroboration, I can agree. But I think that it would be far from '*useful*' to identify the concept of corroboration with that of truth.*[3] This is also avoided in ordinary usage. For one might well say of a theory that it has hardly been corroborated at all so far, or that it is still uncorroborated. But we should not normally say of a theory that it is hardly true at all so far, or that it is still false.

85. *The Path of Science.*

One may discern something like a general direction in the evolution of physics—a direction from theories of a lower level of universality to theories of a higher level. This is usually called the 'inductive' direction; and it might be thought that the fact that physics advances in this 'inductive' direction could be used as an argument in favour of the inductive method.

Yet an advance in the inductive direction does not necessarily consist of a sequence of inductive inferences. Indeed we have shown that it may be explained in quite different terms—in terms of degree of testability and corroborability. For a theory which has been well corroborated can only be superseded by one of a higher level of universality; that is, by a theory which is better testable and which, in addition, *contains* the old, well corroborated theory—or at least a good approximation to it. It may be better, therefore, to describe that trend—the advance towards theories of an ever higher level of universality—as 'quasi-inductive'.

The quasi-inductive process should be envisaged as follows. Theories of some level of universality are proposed, and deductively tested; after that, theories of a higher level of universality are proposed, and in their turn tested with the help of those of the previous levels of universality, and so on. The methods of testing are invariably

*[3] Thus if we were to define 'true' as 'useful' or 'successful' or 'confirmed' or 'corroborated', we should only have to introduce a new 'absolute' and 'timeless' concept to play the role of 'truth'.

based on deductive inferences from the higher to the lower level;*[1] on the other hand, the levels of universality are reached, in the order of time, by proceeding from lower to higher levels.

The question may be raised: 'Why not invent theories of the highest level of universality straight away? Why wait for this quasi-inductive evolution? Is it not perhaps because there is after all an inductive element contained in it?' I do not think so. Again and again suggestions are put forward—conjectures, or theories—of all possible levels of universality. Those theories which are on too high a level of universality, as it were (that is, too far removed from the level reached by the testable science of the day) give rise, perhaps, to a 'metaphysical system'. In this case, even if from this system statements should be deducible (or only semi-deducible, as for example in the case of Spinoza's system), which belong to the prevailing scientific system, there will be no *new* testable statement among them; which means that no crucial experiment can be designed to test the system in question.*[2] If, on the other hand, a crucial experiment can be designed for it, then the system will contain, as a first approximation, some well corroborated theory, and at the same time also something new—and something that can be tested. Thus the system will not, of course, be 'metaphysical'. In this case, the system in question may be looked upon as a new advance in the quasi-inductive evolution of science. This explains why a link with the science of the day is as a rule established only by those theories which are proposed in an attempt to meet the current problem situation; that is, the current difficulties, contradictions, and falsifications. In proposing a solution to these difficulties, these theories may point the way to a crucial experiment.

To obtain a picture or model of this quasi-inductive evolution of science, the various ideas and hypotheses might be visualized as particles suspended in a fluid. Testable science is the precipitation of these particles at the bottom of the vessel: they settle down in layers (of universality). The thickness of the deposit grows with the

*[1] The 'deductive inferences from the higher to the lower level' are, of course, *explanations* (in the sense of section 12); thus the hypotheses on the higher level are *explanatory* with respect to those on the lower level.

*[2] It should be noted that I mean by a crucial experiment one that is designed to refute a theory (if possible) and more especially one which is designed to bring about a decision between two competing theories by refuting (at least) one of them—without, of course, proving the other. (See also note 1 to section 22, and appendix *ix.)

number of these layers, every new layer corresponding to a theory more universal than those beneath it. As the result of this process ideas previously floating in higher metaphysical regions may sometimes be reached by the growth of science, and thus make contact with it, and settle. Examples of such ideas are atomism; the idea of a single physical 'principle' or ultimate element (from which the others derive); the theory of terrestrial motion (opposed by Bacon as fictitious); the age-old corpuscular theory of light; the fluid-theory of electricity (revived as the electron-gas hypothesis of metallic conduction). All these metaphysical concepts and ideas may have helped, even in their early forms, to bring order into man's picture of the world, and in some cases they may even have led to successful predictions. Yet an idea of this kind acquires scientific status only when it is presented in falsifiable form; that is to say, only when it has become possible to decide empirically between it and some rival theory.

My investigation has traced the various consequences of the decisions and conventions—in particular of the criterion of demarcation—adopted at the beginning of this book. Looking back, we may now try to get a last comprehensive glimpse of the picture of science and of scientific discovery which has emerged. (What I have here in mind is not a picture of science as a biological phenomenon, as an instrument of adaptation, or as a roundabout method of production: I have in mind its epistemological aspects.)

Science is not a system of certain, or well-established, statements; nor is it a system which steadily advances towards a state of finality. Our science is not knowledge (*epistēmē*): it can never claim to have attained truth, or even a substitute for it, such as probability.

Yet science has more than mere biological survival value. It is not only a useful instrument. Although it can attain neither truth nor probability, the striving for knowledge and the search for truth are still the strongest motives of scientific discovery.

We do not know: we can only guess. And our guesses are guided by the unscientific, the metaphysical (though biologically explicable) faith in laws, in regularities which we can uncover—discover. Like Bacon, we might describe our own contemporary science—'the method of reasoning which men now ordinarily apply to nature'—as consisting of 'anticipations, rash and premature' and as 'prejudices'.[1]

[1] Bacon, *Novum Organum*, I, 26.

But these marvellously imaginative and bold conjectures or 'anticipations' of ours are carefully and soberly controlled by systematic tests. Once put forward, none of our 'anticipations' are dogmatically upheld. Our method of research is not to defend them, in order to prove how right we were. On the contrary, we try to overthrow them. Using all the weapons of our logical, mathematical, and technical armoury we try to prove that our anticipations were false—in order to put forward, in their stead, new unjustified and unjustifiable anticipations, new 'rash and premature prejudices', as Bacon derisively called them.[*1]

It is possible to interpret the ways of science more prosaically. One might say that progress can '. . . come about only in two ways: by gathering new perceptual experiences, and by better organizing those which are available already'.[2] But this description of scientific progress, although not actually wrong, seems to miss the point. It is too reminiscent of Bacon's induction: too suggestive of his industrious gathering of the 'countless grapes, ripe and in season',[3] from which he expected the wine of science to flow: of his myth of a scientific method that starts from observation and experiment and then proceeds to theories. (This legendary method, by the way, still inspires some of the newer sciences which try to practice it because of the prevalent belief that it is the method of experimental physics.)

The advance of science is not due to the fact that more and more

[*1] Bacon's term 'anticipation' ('*anticipatio*') means almost the same as 'hypothesis' (in my way of using this term). Bacon's view was that, to prepare the mind for the intuition of the true *essence* or *nature* of a thing, it has to be meticulously cleansed of all anticipations, prejudices, and idols. For the source of all error is the impurity of our own minds: Nature itself does not lie. The main function of eliminative induction is (as with Aristotle) to assist the purification of the mind. (*See also* my *Open Society*, chapter 24; note 59 to chapter 10; note 33 to chapter 11, where Aristotle's theory of induction is briefly described.) Purging the mind of prejudices is conceived as a kind of ritual prescribed for the scientist, analogous to the mystic's purification of his soul to prepare it for the vision of God. (*Cf.* section *4 of my *Postscript*.)

[2] P. Frank, *Das Kausalgesetz und seine Grenzen*, (1932). *The view that the progress of science is due to the accumulation of perceptual experiences is still widely held (*cf.* my second Preface, 1958). My denial of this view is closely connected with the rejection of the doctrine that science or knowledge is *bound* to advance, since our experiences are *bound* to accumulate. As against this, I believe that the advance of science depends upon the free competition of thought, and thus upon freedom, and that it must come to an end if freedom is destroyed (though it may well continue for some time in some fields, especially in technology). This view is more fully expounded in my *Poverty of Historicism* (section 32). I also argue there (in the Preface) that the growth of our knowledge is unpredictable by scientific means, and that, as a consequence, the future course of our history is also unpredictable.

[3] Bacon, *Novum Organum* I, 123.

perceptual experiences accumulate in the course of time. Nor is it due to the fact that we are making ever better use of our senses. Out of uninterpreted sense-experiences science cannot be distilled, no matter how industriously we gather and sort them. Bold ideas, unjustified anticipations, and speculative thought, are our only means for interpreting nature: our only organon, our only instrument, for grasping her. And we must hazard them to win our prize. Those among us who are unwilling to expose their ideas to the hazard of refutation do not take part in the scientific game.

Even the careful and sober testing of our ideas by experience is in its turn inspired by ideas: experiment is planned action in which every step is governed by theory. We do not stumble upon our experiences, nor do we let them flow over us like a stream. Rather, we have to be active: we have to '*make*' our experiences. It is we who always formulate the questions to be put to nature; it is we who try again and again to put these question so as to elicit a clear-cut 'yes' or 'no' (for nature does not give an answer unless pressed for it). And in the end, it is again we who give the answer; it is we ourselves who, after severe scrutiny, decide upon the answer to the question which we put to nature—after protracted and earnest attempts to elicit from her an unequivocal 'no'. 'Once and for all', says Weyl,[4] with whom I fully agree, 'I wish to record my unbounded admiration for the work of the experimenter in his struggle to wrest *interpretable facts* from an unyielding Nature who knows so well how to meet our theories with a decisive *No*—or with an inaudible *Yes*.'

The old scientific ideal of *epistēmē*—of absolutely certain, demonstrable knowledge—has proved to be an idol. The demand for scientific objectivity makes it inevitable that every scientific statement must remain *tentative for ever.* It may indeed be corroborated, but every corroboration is relative to other statements which, again, are tentative. Only in our subjective experiences of conviction, in our subjective faith, can we be 'absolutely certain'.[5]

With the idol of certainty (including that of degrees of imperfect certainty or probability) there falls one of the defences of obscurantism which bars the way of scientific advance, checking the boldness of our

[4] Weyl, *Gruppentheorie und Quantenmechanik* (1931), p. 2. English translation by H. P. Robertson: *The Theory of Groups and Quantum Mechanics* (1931), p. xx.

[5] *Cf.* for example note 3 to section 30. This last remark is of course a psychological remark rather than an epistemological one; *cf.* sections 7 and 8.

questions, and endangering the rigour and the integrity of our tests. The wrong view of science betrays itself in the craving to be right; for it is not his *possession* of knowledge, of irrefutable truth, that makes the man of science, but his persistent and recklessly critical *quest* for truth.

Has our attitude, then, to be one of resignation? Have we to say that science can fulfil only its biological task; that it can, at best, merely prove its mettle in practical applications which may corroborate it? Are its intellectual problems insoluble? I do not think so. Science never pursues the illusory aim of making its answers final, or even probable. Its advance is, rather, towards the infinite yet attainable aim of ever discovering new, deeper, and more general problems, and of subjecting its ever tentative answers to ever renewed and ever more rigorous tests.