Logical versus Historical Theories of Confirmation

by ALAN MUSGRAVE

1 Introduction.
2 'Background Knowledge' and the Paradox of Confirmation.
3 The Strictly Temporal View of Background Knowledge.
4 The Heuristic View of Background Knowledge.
5 Background Knowledge or a Background Theory?
6 Requirements for the Growth of Knowledge.
7 Conclusion.

I INTRODUCTION

Thales predicted an eclipse, and became one of the Seven Sages. Since then many have urged that a scientific theory is to be especially prized if it yields successful predictions. For example, Clavius argued that Ptolemaic astronomy must be true because by using its hypotheses 'not only are all the appearances already known accounted for, but also future phenomena are predicted'.\(^1\) Descartes rejected Ptolemaic astronomy, but accepted the methodological principle: we know that our hypotheses are correct, he says, 'only when we see that we can explain in terms of them, not merely the effects we had originally in mind, but also all other phenomena of which we did not previously think'.\(^2\) For Leibniz it was 'the greatest commendation of an hypothesis (next to truth) if by its help predictions can be made even about phenomena not yet tried'.\(^3\) Whewell insisted that good

---

\(^1\) Christopher Clavius, *In Sphaeram Ionnis de Sacro Bosco commentarius* (1602), pp. 518–9; cited from Blake [1960], p. 34. Clavius claimed that the success of Ptolemaic predictions was a powerful argument for the reality of Ptolemy's eccentrics and epicycles and against an instrumentalist interpretation of them as 'geometric fictions': 'For it is incredible that we force the heavens (but we seem to force them, if the Eccentrics and Epicycles are figments, as our adversaries will have it) to obey the figments of our own minds, and to move as we will, or in accordance with our principles.' Later it was successful predictions which led Duhem to water down his instrumentalism with the idea of 'natural classification'.

\(^2\) René Descartes, *Principles of Philosophy* (1644), Part III, xlii.

\(^3\) Leibniz, Letter to Conring (1678); cited from Lakatos [1970], p. 123.
hypotheses ‘ought to foretell phenomena which have not yet been observed’. Duhem agreed.

Quotations like this could be multiplied, from both scientists and philosophers of science. The intuitions they express are that a good hypothesis must not only account for known phenomena, but must successfully predict new ones; that explaining known facts is one thing, predicting new facts another; and that in assessing the confirmation or evidential support of a hypothesis, we must take into account especially (perhaps even exclusively) the success or failure of its predictions.

These intuitions have usually received a dusty response from philosophers interested in the theoretical analysis of evidential support or confirmation. According to modern logical empiricist orthodoxy, in deciding whether hypothesis \( h \) is confirmed by evidence \( e \), and how well it is confirmed, we must consider only the statements \( h \) and \( e \), and the logical relations between them. It is quite irrelevant whether \( e \) was known first and \( h \) proposed to explain it, or whether \( e \) resulted from testing predictions drawn from \( h \). I shall call this cornerstone of logical empiricism the purely logical (or logical, for short) approach to confirmation.

We find it in Mill, who was amazed at Whewell’s view that ‘an hypothesis . . . is entitled to a more favourable reception, if besides accounting for all the facts previously known, it has led to the anticipation and prediction of others which experience afterwards verified. Such predictions and their fulfilment are, indeed, well calculated to impress the uninformed . . . But it is strange that any considerable stress should be laid upon such a coincidence by persons of scientific attainments’. Keynes agreed, and insisted that the ‘peculiar value of prediction . . . is altogether imaginary . . . The question as to whether a particular hypothesis happens to be propounded before or after examination of [its instances] is quite irrelevant’. Hempel tries to answer the qualitative question of whether or not \( e \) confirms \( h \) by considering only the logical forms of \( e \) and \( h \). And Carnap’s quantitative degrees of confirmation are also to depend entirely upon the

---

2 See section 2 of chapter 5 of Duhem’s [1954], entitled ‘Theory anticipating experiment’.
3 J. S. Mill, System of Logic, III, xiv, 6. Cf. Medawar [1967], p. 141, and Lakatos [1970], p. 123. Mill’s main point (that successful prediction cannot conclusively prove the theory involved) was, of course, quite right—and Clavius, Descartes and Whewell were wrong.
5 Hempel says: ‘. . . it seems reasonable to require that the criteria of empirical confirmation . . . should contain no reference to the specific subject matter of the hypothesis or of the evidence in question; it ought to be possible, one feels, to set up purely formal criteria of confirmation in a manner similar to that in which deductive logic provides purely formal criteria for the validity of deductive inference’ (Hempel [1965], p. 10).
logical forms of the two sentences concerned. On this view, so far as problems of confirmation or evidential support are concerned 'it is quite irrelevant whether in fact scientists always or usually or never make their observations before conceiving their theories or vice versa'.

So we see that those who follow the logical approach to confirmation show scant respect for the intuition that successful prediction matters. After all, they argue, scientists use evidence to evaluate theories by checking whether their consequences are true. If the evidence shows that some consequence of a theory is true, then this cannot depend upon whether the evidence came to be known before the theory was proposed or afterwards. Such historical considerations, interesting though they may be, ought not to affect questions of confirmation or evidential support. For these questions are objective ones, to be settled solely by inspecting the logical relationship between theory and evidence.

Yet the purely logical approach to confirmation has run into some notorious difficulties—the most notorious of all being the paradox of confirmation. And the best solution to that paradox, or so it seems to me, points towards a different approach to confirmation, and one which does more justice to the intuition that successful predictions are important. It does so, as Lakatos was the first to emphasise, by introducing an historical ingredient into confirmation. Hence I shall call it the partly historical or logico-historical (or historical, for short) approach to confirmation.

The main aim of this paper is to examine the acceptability of this logico-historical approach to confirmation. I shall begin with the paradox of confirmation, and with what I regard as the best solution to it, which was developed by Watkins using some of Popper's ideas. I will then distinguish three different ways of introducing an historical ingredient into confirmation, and assess their merits. I shall ask, in particular, whether once we move away from the logical approach, we can avoid introducing an undesirable subjectivity and arbitrariness into questions of evidential support—whether we can avoid saying that the evidential support of a theory depends upon personal or psychological facts about the man who proposed that theory.

2 'BACKGROUND KNOWLEDGE' AND THE PARADOX OF CONFIRMATION

To set the scene, let us restate the paradox of confirmation, and see how Watkins solves it by introducing the notion of 'background knowledge'. The paradox arises if we accept the following two principles:

(I) Any positive instance of a hypothesis confirms that hypothesis;
(E) If e confirms \( H_1 \), and \( H_2 \) is logically equivalent to \( H_1 \), then e also confirms \( H_2 \).

As is well known, these two intuitively obvious principles lead to the intuitively absurd result that any object which is not a counter-instance to a universal hypothesis is a confirming instance to it. Principles (I) and (E) together render confirmation ridiculously easy. To use the usual example, any statement about the colour of any object will, provided it is not a report of a non-black raven, confirm 'All ravens are black'. As Goodman remarked, 'the prospect of being able to investigate ornithological theories without going out in the rain is so attractive that we know there must be a catch in it'.

Solutions to the difficulty are legion, and I shall not attempt to canvas them all. Hempel's own solution to the paradox is simple—we must, he says, simply get used to it. According to Hempel, 'the impression of a paradoxical situation is not objectively founded; it is a psychological illusion'. The illusion is fostered by two mistakes. First, we think that a statement like 'All ravens are black' is only about ravens, whereas in fact it asserts something about every object in the universe (that it is either not a raven or black). Second, we allow additional information to intrude. We think it odd that observations of objects inside a room can confirm 'All ravens are black' because we know in advance that there are no ravens inside rooms, and hence no non-black ones. But to take account of this additional knowledge is, says Hempel, to 'fail to observe the methodological fiction, characteristic of every case of confirmation, that we have no relevant evidence for \( H \) other than that included in \( E \). He says we must be 'careful to avoid this tacit reference to additional knowledge (which entirely changes the character of the problem)'.

Watkins solves the paradox by recognising Hempel's 'methodological fiction' for what it is, namely, a fiction and one which begs the question in favour of the purely logical approach. Since the scientist is never in the position of having no information other than \( E \), Watkins tries to take his additional information (his 'background knowledge') into account. And this does, as Hempel said it would, entirely change the character of the problem.

1 Whether e is a 'positive instance' of \( h \) is supposed to depend on the logical forms of e and \( h \). Intuitively, an object which satisfies both the antecedent and consequent of a universal law will be described in a positive instance of that law. Hempel's own version of (I) reads: e confirms \( h \) if e logically implies the 'development' of \( h \) for the individuals mentioned (essentially) in e. For details see Hempel [1965], pp. 35–9.
2 Goodman [1954], pp. 71–2. 3 Hempel [1965], p. 18. 4 Hempel [1965], pp. 19–20. 5 Watkins' solution is contained in his [1964]. Popper and Watkins prefer to speak of 'corroboration' rather than 'confirmation' in order to distinguish their view from the purely logical approach. I shall not follow them in this.
Watkins begins with Popper's thesis that genuine confirmations can result only from the failure of genuine attempts to refute a hypothesis, from severe tests. But how are we to recognise a 'genuine attempt to refute' or a 'severe test'? Popper occasionally suggests that it is a psychological affair: for a test to be severe the experimenter who performs it must be sincerely trying to overthrow the theory tested. I have criticised this view elsewhere, and will say no more about it here except that the severity of tests must be an objective or impersonal matter, rather than a subjective or personal one. Elsewhere Popper says the same, as when he claims that the severity of tests can be objectively compared once we take 'background knowledge' into account.

According to this objective theory, the severest tests of a hypothesis are those which in the light of our background knowledge are most likely to refute it. And if background knowledge tells us that some possible observation or experiment will not refute our hypothesis, then making that observation or experiment will not be a severe test and so cannot result in a confirmation. Applied to the paradox of the ravens, suppose background knowledge tells us that ravens are not to be found in rooms. Background knowledge therefore entails, about any particular object in a room, that it will not be a non-black raven—which is exactly what 'All ravens are black' predicts about that object. Thus background knowledge tells us that the prediction from our hypothesis is true, that we are not going to refute our hypothesis by observing objects in rooms. Since such an observation would not represent a severe test, it cannot result in a confirmation of 'All ravens are black'.

Suppose, on the other hand, that background knowledge tells us that the prediction from our hypothesis is likely to be false. To test such a prediction will be a severe test, and if the hypothesis passes the test it will be genuinely confirmed by the resulting evidence statement. The severest tests of all will be where background knowledge logically implies the negation of the prediction. A test of such a prediction will be a crucial test between the hypothesis and background knowledge.

Watkins makes it clear that if we assume that we have no background knowledge at all, then we are back in the Hempelian situation where all tests are equally severe (or rather, equally lacking in severity). Watkins also makes it clear that the introduction of (non-empty) background knowledge avoids the paradox by denying that every positive instance of a hypothesis

---

1 See section 3 of my [1974].
2 See, for example, Popper's [1963], p. 388.
3 It almost goes without saying that our 'background knowledge' may in fact be false, and that taking it into account may lead us astray. But with nothing to guide him, the scientist would have no clue as to which of the infinitely many possible tests he should perform.
confirms it. Finally, Watkins shows that sometimes so-called ‘irrelevant instances’ (evidence concerning non-ravens) may confirm ‘All ravens are black’.\(^1\) Anyone who finds it too counterintuitive to suppose that a positive instance need not confirm, while an irrelevant instance might, will have to reject Watkins’ solution.

Watkins’ solution also contains the principle that a hypothesis cannot be confirmed by the facts it was devised to explain. We need only assume that the facts a hypothesis is devised to explain are contained in the ‘background knowledge’ to that hypothesis. Clearly such facts cannot represent the results of severe tests, and so cannot confirm the hypothesis. It follows that a hypothesis can be confirmed only by testing new independent predictions. The intuitive idea that successful prediction is important is given a very sharp formulation: so far as evidential support is concerned, successful prediction is all-important.

This last point is contained in Popper’s paper *The Aim of Science* published in 1957. Popper had earlier maintained that to be scientific a hypothesis must be empirically testable, it must exclude some observable state of affairs. But any hypothesis which explains the observed facts \(e\) is testable in this sense—for it excludes the observable state of affairs in which \(e\) is not the case. Thus Popper’s testability requirement does not exclude explanatory systems which, though they yield known facts, do not lead to any new predictions. To avoid such explanations (ad hoc explanations, as he calls them), Popper strengthened his requirement of testability to a requirement of independent testability. Generations of scientists have been suspicious of ad hoc explanations. And Popper gives their suspicion a very strong rationale: only severe tests can confirm, only independent tests can be severe, so an *ad hoc* explanation (an explanation which is not independently testable) is unsatisfactory simply because it cannot genuinely be confirmed by empirical evidence.\(^2\)

---

\(^1\) Convincing examples of this are hard to find for trivial hypotheses like ‘All ravens are black’. But suppose background knowledge tells us that all birds which have a certain property \(P\) are ravens. And suppose an observer spots a white bird with property \(P\), call it \(a\). Background knowledge together with \(Pa\) predicts that \(a\) is a raven—while ‘All ravens are black’ together with ‘\(a\) is not black’ predicts that \(a\) is not a raven. Inspection of \(a\) to see if it is a raven will be a crucial test between the hypothesis and background knowledge, which might confirm the hypothesis. Yet if it does confirm it, the observation report will be of the form ‘\(a\) is a non-black non-raven’. Agassi has given a more plausible example of the same kind: cf. the account in Watkins [1964], pp. 108–9.

\(^2\) It was Lakatos (in his [1968], pp. 376–8) who first emphasised the importance of the transition from testability to independent testability. As he says, it is independent testability which is important if our main interest is in the growth of knowledge.

Popper objected to *ad hoc* hypotheses already in his [1934], pp. 82–3, but there he used the term ‘*ad hoc*’ in a more specialised sense. He pointed out that it is always possible to evade a refutation of a theoretical system by introducing auxiliary assumptions into that system. It would be unreasonable to exclude *all* such auxiliary assumptions, since some of them led to great theoretical triumphs (e.g. the discovery of Neptune). Therefore
Popper summarised these ideas in formulas, expressed in terms of logical probabilities, for the severity of tests and the degree of confirmation. The main idea is that the severity of a test resulting in evidence \( e \) on the hypothesis \( h \) in the light of background knowledge \( b \) is greater the more probable is \( e \) given \( h \) and \( b \) and the less probable is \( e \) given \( b \) alone: \( S(e, h, b) = p(e, h \& b) - p(e, b) \). The degree of confirmation of \( h \) by \( e \) in the presence of \( b \) is defined so that it is proportional to the severity of the test. Thus if the severity of the test is zero, so is the degree of confirmation resulting from it.

Now so far the notion of ‘background knowledge’ has been left rather vague. And as we will see, different variants of the historical approach to confirmation result from construing background knowledge in slightly different ways. But before proceeding, I should explain why I call a theory which takes into account background knowledge (however construed) a partly historical theory. I do so because once we have decided what sort of thing background knowledge is to contain, it will presumably be a historical task to determine its actual contents in a given case. Thus all variants of the historical approach will make the confirmation of a scientific theory somehow depend upon the historical setting in which that theory was proposed. Of course, once the actual content of background knowledge has been ascertained by historical investigation, the analysis of confirmation proceeds logically. But we investigate the logical relations between three things (theory, evidence, and background knowledge) and not two as in the purely logical approach.

How, then, is ‘background knowledge’ to be construed—what sort of thing does it contain? As soon as we pose this question, the spectre of subjectivism and relativism rises before us. For suppose we construe the

---

1 See Popper [1963], Appendix 2. Popper’s formulas contain Mackie’s ‘inverse principle’ or ‘relevance criterion of confirmation’; cf. Mackie [1969]. Popper himself does not think his formulas can yield numerical degrees of confirmation. For them to do so, we would have to be able to assign numerical values to the function \( p(a, b) \)—and Popper does not think this can be done in all cases.

Moreover, these formulas do not do justice to the importance of crucial tests between a new and an old hypothesis. For background knowledge has to be pruned so that the conjunction \( 'a \& b' \) is consistent, if the formulas are not to yield absurd results. For this reason, Lakatos modifies Popper’s formulas: cf. Lakatos [1968], p. 375, note 2, and below, p. 17.
background knowledge to an hypothesis in a subjective or personal sense—as the collection of facts and theories which a particular scientist actually accepts as unproblematic while he considers that hypothesis. There is every reason to believe that different scientists will have different 'background knowledges' in this sense. And it will follow from the historical theory, thus construed, that the evidential support of a theory will depend upon which scientist we take to be considering that theory.\(^1\)

The attempt to avoid this relativistic conclusion leads us to the first variant of the historical theory which I shall consider.

3 **THE STRICTLY TEMPORAL VIEW OF BACKGROUND KNOWLEDGE**

Suppose we insist that the 'background knowledge' to a given theory contains all the relevant experimental results, hypotheses, etc., which are 'known to science' when that theory is proposed. The fact that different individuals will actually be aware of different portions of this background knowledge will cease to affect the confirmation of the new theory. Instead, we will have a *strictly temporal* view of background knowledge, and hence of independent testability. Facts 'known to science' before a hypothesis is proposed will not be able to confirm that hypothesis, since they will already be contained in background knowledge and cannot represent the results of severe tests. On this view, the only tests that can result in confirmations are tests of predictions of *new effects*. And a theory is independently testable only if it predicts a *novel fact*, a fact not 'known to science' when the theory was proposed.\(^2\)

This view presupposes, of course, that the historian can settle the question of when a particular statement becomes 'known to science'. And it must be admitted that a certain penumbra of vagueness may surround this matter. Does a statement become 'known to science' when some scientist first thinks of it?—or when he first writes it down?—or when it is first

---

\(^1\) Stove voiced essentially this objection to introducing background knowledge into confirmation in 1960: 'For consider two men, both of whom now defend a certain theory by appealing to certain observations, the theory and the observations being exactly the same in the two cases; but one of whom made the observations after, and because of, having conceived the theory, while the other conceived the theory after, and because of, having made the observations. Here, clearly, we must say, either that the corroboration, if any, which the observations afford is exactly the same in the two cases, or that the mental history of an investigator necessarily enters into the question of the evidential weight of his observations' ([1960], pp. 178–9).

\(^2\) According to Zahar ([1973], pp. 101–2), Lakatos was the first to adopt this strictly temporal view explicitly—and as we shall see, (below, p. 11), Lakatos certainly does try to soften some of the harsher consequences of such a view. Despite this, however, I have not been able to find where Lakatos explicitly adopts the strictly temporal view. Whenever he actually defines what he means by independent testability (or 'boldness', or the 'novelty of a fact') he adopts a different view, the one to be considered in section 5 below (I cite his definitions on p. 16, note 1).
published—or perhaps when it is abstracted in a leading journal? Often it will not matter how we answer, for a very loose dating will suffice. And where it does matter, historians are quite used to settling such questions.\footnote{Zahar ([1973], p. 103, note 1) reports an amusing argument, due to John Worrall, which highlights the difficulties. Suppose scientist A devises theory $T$ to account for the known facts $e$, and then uses $T$ to predict the new fact $f$. A gets experimentalist B to check his prediction, and he confirms it. They send off their papers to different journals. But the theoretical journal takes longer to publish, so that the experimental confirmation of $f$ appears in print before the theoretical prediction of it. On a strictly temporal view, can $f$ confirm $T$? It can, of course. First of all, the two journals will probably print the dates of receipt of their articles, enabling $T$ to get credit for having predicted $f$. Even if they do not (or if A’s manuscript was delayed in the post), $B$ may refer to $A$’s theoretical prediction in his paper. Even if he does not, he will probably admit if asked that he performed his experiment as a result of hearing about $A$’s prediction. And if he does not, $A$ may be able to produce evidence (from manuscripts, letters, diaries) that his theory did predict $f$. In other words, the strictly temporal theorist must, in hard cases like this, adopt rather sophisticated historiographical methods.}

Given that we can make sense of the idea of a given body of knowledge in a certain field at a particular time, the strictly temporal view does enable us to avoid introducing extreme relativism into confirmation.

But although an extreme relativity to persons is avoided, a less extreme though quite sharp historical relativity is introduced. Following Lakatos, I shall explain what this means by contrasting two schematic situations:\footnote{Cp. the schematic situations described by Lakatos in his [1970], pp. 151–2. Incidentally, in my diagrams time flows from left to right, downward arrows pointing left represent explanation, downward arrows pointing right represent successful prediction, and upward arrows represent refutation.}

\textit{Situation (1)}: Phenomenon $e_1$ is known, and theory $T$ proposed which explains it. $T$ is independently testable, since it predicts the unknown phenomena $e_2, \ldots, e_n$. These bold predictions are checked and confirmed. We have a marvellous success story.

\begin{center}
\begin{tikzpicture}
  \node (e1) at (0,0) {$e_1$};
  \node (e2) at (1,0) {$e_2$};
  \node (en) at (2,0) {$\ldots$};
  \node (en1) at (3,0) {$e_n$};
  \node (t) at (1.5,1) {$T$};

  \draw [->] (t) to (e1);
  \draw [->] (t) to (e2);
  \draw [->] (t) to (en);
  \draw [->] (t) to (en1);
\end{tikzpicture}
\end{center}

\textit{Situation (2)}: The same phenomena $e_1, e_2, \ldots, e_n$ are all known in advance, and the same theory $T$ proposed which explains them all. But now $T$ is not independently testable, since it predicts no new phenomena. The historical theorist regretfully concludes that $T$ is \textit{ad hoc}.

\begin{center}
\begin{tikzpicture}
  \node (e1) at (0,0) {$e_1$};
  \node (e2) at (1,0) {$e_2$};
  \node (en) at (2,0) {$\ldots$};
  \node (en1) at (3,0) {$e_n$};
  \node (t) at (1.5,1) {$T$};

  \draw [->] (t) to (e1);
  \draw [->] (t) to (e2);
  \draw [->] (t) to (en);
  \draw [->] (t) to (en1);
\end{tikzpicture}
\end{center}
Now in both these situations we have exactly the same theory \( T \) and exactly the same body of evidence—the difference lies only in the order in which theory and evidence were discovered. For the logical theorist \( T \) will be equally well confirmed in both cases. But for the historical theorist \( T \) will be marvellously confirmed in the first situation, and cannot be confirmed at all in the second.\(^1\)

It might be objected that my situation (2) never actually arises. For empirical facts are never simply accumulated as that situation supposes, but rather are only discovered by testing predictions from antecedently-held hypotheses. This interesting Popperian conjecture may well be true—I happen to think that, by and large, it is true.\(^2\) But even if we accept it, we cannot avoid the historical relativity of confirmation. For we can replace situation (2) by:

**Situation (3):** Phenomenon \( e_1 \), and theory \( T_1 \) proposed which explains it. \( T_1 \) is independently testable and forbids phenomenon \( e_2 \). This new prediction is checked and found to be false. \( T_1 \) is rejected, and \( T_2 \) proposed which explains both \( e_1 \) and \( e_2 \). \( T_2 \) is also independently testable, and forbids \( e_3 \). Again the new prediction is checked and found to be false. The process continues until \( e_1, \ldots, e_n \) are known, and \( T \) proposed which explains them all. But \( T \) is not independently testable, since it predicts no new phenomena. As in situation (2), the historical theorist regretfully concludes that \( T \) is *ad hoc*.

Here all the empirical facts (except for \( e_1 \)) are discovered by testing predictions from antecedently held theories, precisely as the Popperian historical conjecture requires. Yet just as before, the historical theorist will

---

1 The inductivist regards theories which anticipate facts as premature speculation. So the inductivist teacher of science will present the history of theory \( T \) as if situation (2) were the case. But if he has Popperian students, he will only succeed in convincing them that \( T \) is *ad hoc*. If science actually proceeds, by and large, as in situation (3), then its inductivist virtues will be apparent only if it is taught unhistorically, while its Popperian virtues will be apparent only if it is taught historically.

Popper makes the interesting suggestion that the conventionalist view of physical theory arises from taking an unhistorical view of its development—see his [1963], p. 240.

2 But *cp. below*, note 1 to page 18.
have to say that $T$ is marvellously confirmed in situation $(x)$, and not con-
firmable at all in situation $(z)$.\footnote{In terms of Popper's three requirements for the growth of knowledge (to be discussed in section 6), in situation $(z)$ each of the $T_i \ (1 < i < n - 1)$ satisfies the second requirement (independent testability) but fails the third (independent confirmation), while $T$ fails to surmount even the second hurdle. Yet so far as the evidence goes, $T$ may be true while all its predecessors are false!}

Those whose intuitions favour a purely logical approach to confirmation will, no doubt, find this kind of historical relativity absurd. But the purely logical approach has some absurd results of its own. It seems that in this area some of our intuitions have to go—and I for one am prepared to accept that the historical background to a theory might make a difference to its confirmation.

But ought strictly temporal considerations count for so much as this variant of the historical theory suggests? The strictly temporal view of independent testability is certainly a very stringent one. It means that Galileo's and Kepler's laws (or facts about the tides, or the precession of the equinoxes) cannot confirm Newton's theory, simply because they were 'known to science' before that theory was proposed.\footnote{In fact, of course, Newton predicted deviations from Galileo's and Kepler's laws, most of which represented novel facts. Since Newton's theory was devised to explain Galileo's and Kepler's laws, it is the paradigm of a non-$\textit{ad hoc}$ theory, since it corrects its own $\textit{explanandum}$. See Popper's [1972], p. 202.} It means that the Michelson-Morley experiment cannot confirm the Special Theory of Relativity, nor can the anomalous precession of Mercury's perihelion confirm the General Theory. And since they too were known in advance, Balmer's empirical formulas for the emission spectrum of excited hydrogen cannot confirm Bohr's theory of the hydrogen atom.

In connection with this last example, Lakatos tries to water down the strictly temporal view of background knowledge. He suggests that if a known fact is 'reinterpreted' in the light of a new theory, it can turn into a novel fact for that theory: 'we should certainly regard a newly interpreted fact as a new fact, ignoring the insolent priority claims of amateur fact collectors'.\footnote{Lakatos [1970], p. 157; the whole sentence is italicised in the original.}

Now it will hardly do, especially from a methodologist who professes to take the judgements of scientists seriously, to consign a great experimentalist like Balmer to the flames as an 'amateur fact collector'. Moreover, as Zahar points out, Lakatos's watering-down of the notion of a novel fact actually obliterates it altogether.\footnote{See Zahar [1973], p. 102. Noretta Koertge generously chose to ignore what she called Lakatos's 'lapse into the we-live-in-a-different-world-after-a-revolution syndrome', and insisted that if he 'wants to count a research-programme as progressive when it is found to predict a known result, he should just say so'; see her [1971], p. 171, note 5.} Any deduction of an old fact from a new
theory can be said to involve a ‘reinterpretation’ of that fact. And now any *ad hoc* theory can claim evidential support from the old, but ‘newly interpreted’, facts which it explains.

Zahar rightly deplores this wholesale retreat—yet he still wants the Michelson–Morley experiment to confirm the theory of relativity. He therefore proposes a new and ingenious view of independent testability or of the novelty of a prediction. This brings me to the second variant of the historical theory.

4 THE HEURISTIC VIEW OF BACKGROUND KNOWLEDGE

Zahar suggests that the novelty of a fact should not depend merely on temporal considerations. Instead, an old fact can be novel with respect to a new theory provided that theory was not devised to explain it. Hence temporal novelty is a sufficient condition for novelty in Zahar’s sense, but not a necessary one. If the Michelson–Morley experiment played no role in the construction of the Special theory, then it can still confirm that theory (and similarly for the other examples mentioned above). Or, to revert to our schematic situations, \( T \) is no longer automatically *ad hoc* in situations (2) and (3). It all depends upon whether the scientist who devised \( T \) made use of all the known facts \( e, \ldots, e_n \) in doing so—if he did \( T \) is *ad hoc*, if he did not it is independently testable.

Zahar’s proposal involves a *heuristic view of background knowledge*. The background knowledge to a new theory does not contain all the previously known facts which it explains, but rather only the previously known facts which it was designed to explain and which played a heuristic role in its construction.

Now Zahar is quite right to insist that the facts which a new theory is *devised* to explain are likely to be less extensive than the previously known facts which it does actually explain. There is even a general argument why this is so: that one or more apparently unrelated facts are connected with each other often becomes clear only in the light of some theory—and so it is unlikely that that theory was devised to account for all of them. Thus we must distinguish the actual *explanandum* for a theory from the previously known facts which it explains.

But the novelty of Zahar’s methodological proposal is that this distinction should have repercussions for the new theory’s evidential support. According to Zahar, the way in which a scientist happens to have arrived at his theory can affect the extent to which that theory is confirmed by empirical evidence. To assess the evidential support of a theory ‘one has to take into account the way [it] is built and the problems it was designed to solve’.¹

¹ Zahar [1973], p. 103 (italicised in the original).
Thus confirmation depends, not on comparatively public historical matters like the dates of discoveries, but on private biographical details about the man who proposed the theory in question. Zahar is well aware that his view 'implies that the traditional methods of historical research are even more vital for evaluating experimental support than [anybody] had previously suggested. The historian has to read the private correspondence of the scientist whose ideas he is studying . . . to disentangle the heuristic reasoning which the latter used in order to arrive at a new theory'. ¹ According to Zahar, then, the private correspondence, diaries, manuscripts, etc., of an individual scientist are no longer merely of historical or heuristic interest—their contents may also help to determine the evidential support of his theory.

Now the first thing to be said about this proposal is that it will make the evidential support of a theory rather difficult to determine. Scientists are notoriously reticent about the routes to their discoveries. Most of them assume that the way in which a theory is arrived at is irrelevant to its value, including its evidential support. Zahar disputes this assumption (and the distinction between the 'context of discovery' and the 'context of justification' which philosophers of science have erected upon it). A scientist's own account of his discovery, often written when he is too old to do anything better than reminisce about 'how I did it', will often be distorted by philosophical preconceptions (including, perhaps, Zahar's should it gain currency amongst scientists). But Zahar can, of course, reply to this that the historian is used to dealing with such difficulties.

A more important question is whether, by departing from the austere standards of the strictly temporal view, Zahar can avoid making the evidential support of a theory a person-relative affair. Suppose two scientists, A and B, independently and at about the same time, propose the hypothesis h. And suppose that A devised h to account for the known facts e₁ and e₂—while B, less au fait with current literature, devised h only to account for e₁. Does Zahar's view entail that e₂ confirms h as proposed by B, but does not confirm h as proposed by A? The historian might salute B's ingenuity (or, perhaps, remark upon his good luck). But will Zahar say that B's ignorance of the literature means that h as proposed by him has more empirical support than the same hypothesis as proposed by A?

I suspect that Zahar would wish to avoid conclusions like this one. For he repeatedly insists that disentangling the heuristic route to a discovery is not, as he calls it, a 'psychologicist' affair.² Thus, he might say, a scientist can

¹ Zahar [1973], pp. 103–4.
² Thus, for example, Zahar insists that our purpose in reading the correspondence of the scientist who invented the theory we are interested in 'will not be to delve into the psyche of the scientist, but to disentangle the heuristic reasoning which the latter used in order
be aware of some extant proposition without it playing a heuristic role in his theorizing (can he, I wonder, be unaware of a proposition which does play a heuristic role?). For example, even if Einstein was aware of the results of the Michelson–Morley experiments, this does not mean that they played any role in the construction of his theories.

But how successful are these disclaimers? Is disentangling a process of heuristic reasoning really so ‘non-psychological’ as Zahar wants it to be? Zahar provides us with a fascinating example—the heuristic route to Einstein’s theories of relativity. Now I shall not dispute the accuracy of Zahar’s account, for I am hardly competent to do so. But I do claim that, unless it is a mere fairy-story with no pretensions to historical accuracy, it must contain historical assertions about the actual beliefs, principles used, etc., of a particular person.

For example, according to Zahar two metaphysical principles and two empirical facts played an important heuristic role in the genesis of the theory of relativity. The metaphysical principles are: Nature is simple or coherent (its associated methodological prescription says that adequate theories should be simple or coherent); and there are no accidents in Nature (its associated methodological prescription is that observed symmetries must be matched by theoretical symmetries). The empirical facts are the results of electromagnetic induction experiments, and the equality of inertial and gravitational mass.¹ Now Zahar must claim, it seems to me, that Einstein actually made use of the metaphysical principles, and actually considered the empirical facts. And such claims are psychological or personal claims about a particular historical figure. But if the evidential support of a theory depends upon the heuristic route to it, and the heuristic route to it involves facts about the man who took that route, I cannot see how Zahar can avoid the conclusion that the evidential support of a theory depends upon facts about the man who proposed it. If different scientists take different heuristic routes to the same theory, then the evidential support of that theory as proposed by one of them might be different from its evidential support as proposed by the other. In short, Zahar’s view makes confirmation a person-relative affair.

to arrive at a new theory’ ([1973], p. 104). Later, when he comes to his case-study, Zahar begins by asking why Einstein objected to classical physics—and he takes pains to reassure us that ‘the answer to this question will not be a psycholgistic answer; I shall not for example be indulging in speculations about Einstein’s childhood’ ([1973], p. 223).

¹ See Zahar ([1973], pp. 223–7. Incidentally, Einstein’s heuristic (as reconstructed by Zahar) bears little resemblance to the ‘positive and negative heuristic’ of Lakatos’s methodology of scientific research programmes. Einstein’s heuristics are closer to the sort of ‘influential metaphysics’ from which Lakatos disassociates himself (cf. his [1970], pp. 183–4). This obviously bears upon Zahar’s claim that he is presenting a case-study of the ‘methodology of scientific research programmes’.
At this point Zahar might be tempted to invoke Lakatos's notion of rational reconstruction. He might say that he is not interested in whether Einstein actually considered the empirical facts he mentions in the light of the metaphysical principles he formulates. He is only 'rationally reconstructing' a heuristic which Einstein might have, or perhaps ought to have, followed. Kuhn has already objected that Lakatos's rational reconstructions are not so much history as philosophy fabricating examples.¹ More importantly, since philosopher-historians of science can provide different rational reconstructions of the genesis of the same theory, the evidential support of a theory will now depend upon which philosopher-historian of science we consider. I doubt that scientists themselves will be prepared to hand over to philosophers of science the task of assessing the evidential support of their theories.

This is not to deny that the notion of 'rational reconstruction' has a part to play in the history of science. As Mark Twain said, in history nothing ever happens in the right place or at the right time—and it is the job of the historian to remedy this defect. The historian can, and does, rationally reconstruct an episode and then compares his reconstruction with the actual history—and the exercise can be most illuminating. But when a historian of science rationally reconstructs the problem-situation for a new theory, surely the most sensible reconstruction will consist of the known problems actually solved by the theory in question. And this will lead us back to a strictly temporal view of 'novelty' or of independent testability.

But if the strictly temporal view is too austere, and Zahar's heuristic view too relativistic, can we preserve any of the insights of the historical approach to confirmation? I think that we can—which brings me to the third variant of the historical approach.

5 BACKGROUND KNOWLEDGE OR A BACKGROUND THEORY?

This third variant starts from the basic idea that scientists operate with competing theories, and that they use empirical evidence to try to decide between them. The basic question is not so much 'Does evidence e confirm hypothesis h?' but rather 'Does evidence e support h₁ more than h₂?'. This suggests that in assessing the evidential support of a new theory we should compare it, not with 'background knowledge' in general, but with the old theory which it challenges. Lakatos, who developed this third variant, calls the old theory the 'background theory' or the 'touchstone theory'.²

According to this view, a new theory is independently testable (or predicts a 'novel fact') if it predicts something which is not also predicted

¹ Kuhn [1971], p. 143.
² Lakatos [1968], pp. 375–90.
by its background theory.\footnote{Or as Lakatos put it, the new theory should have potential falsifiers which are not also potential falsifiers of the old theory ([1968], pp. 376–7). The same view is expressed differently in his [1970], p. 116: to be independently testable or have ‘excess empirical content’ over its predecessor a new theory must predict ‘novel facts, that is, facts improbable in the light of, or even forbidden by’ the old theory. Clearly, a ‘novel fact’ thus defined may very well be a known fact—and hence Lakatos does not explicitly adopt a strictly temporal view of novelty (but see the next footnote).} Hence there are two kinds of independent or novel predictions, tests of which are severe and might result in confirmations. First, there are predictions which conflict with the predictions of the background theory—tests of these will be crucial tests between the new theory and the old.\footnote{At one point Lakatos says that where the old theory $T_1$ has been refuted by evidence $e$ we should make the conjunction $T_1 \land e$ the background theory for any new theory. This proposal would take him close to a strictly temporal view of independent testability. But the proposal is unsatisfactory. It means that crucial tests between the old theory and the new need not be severe ones. Indeed, since $T_1 \land e$ is, by assumption, inconsistent, and since inconsistent premises imply anything, the proposal means that with such a ‘background theory’ no test can be severe. The proposal was, therefore, a slip and I shall say no more about it (cf. Lakatos [1968], p. 375, last paragraph).} Second, there are predictions concerning phenomena about which the background theory predicts nothing at all—tests of these will also be independent tests and severe ones (in the light of ‘background knowledge’ construed as a competing theory).

This third variant gives different results than the other two variants. First, it only partially endorses the view that a theory cannot be confirmed by facts contained in its explanandum. For suppose that the (rationally reconstructed) explanandum for a new theory contains the successes and failures of the old theory. The ‘failures’ being known empirical facts which contradict or cannot be explained by the old theory. According to our third variant, such facts can confirm the new theory despite being known before that theory was proposed. Facts which contradict the old theory and are predicted by the new represent the results of crucial tests between them—and so, it seems, a crucial test between $T_1$ and $T_2$ can be performed even before $T_2$ is proposed.\footnote{This should cause no surprise. Of course, before $T_2$ is proposed such an experiment will not appear to be crucial. As Lakatos (and others) have emphasised, before the new theory is available refutations of the old are classed as ‘unexplained anomalies’—their crucial character only becomes apparent with hindsight (see, for example, Lakatos [1970], pp. 154–9).}

So we see that this third variant does not involve a strictly temporal view of the ‘novelty’ of a prediction. A known fact can confirm a new theory if it contradicts, or cannot be explained by, its predecessor. And a fact discovered by testing a prediction from the new theory will not confirm it if it can also be predicted by the old theory. Thus temporal novelty of a factual discovery is neither sufficient for it to confirm a theory, nor is it necessary.

Our third variant might agree with some of Zahar’s conclusions, but it gives different reasons for them. The Michelson–Morley experiment can
confirm the Special Theory of Relativity, and the Mercury perihelion the General Theory, because they refuted the previous theories. That these experimental results played no role in the construction of Relativity Theory may be of great heuristic interest—but according to our third variant it is quite irrelevant to questions of evidential support.

To see that the third variant will give different results from Zahar's, imagine the following situation. The existing theory $T_1$ explains $e_1$ and $e_2$ but cannot explain $e_3$. A new theory $T_2$ is proposed and $e_1$ and $e_3$ play a heuristic role in the construction of it, while $e_2$ does not. Now according to the third variant $e_3$ will confirm $T_2$, while according to Zahar it will not. And according to the third variant $e_2$ will not confirm $T_2$, while according to Zahar it will. (On a purely historical view, none of $e_1$, $e_2$, $e_3$ can confirm $T_2$.) As we can see from this, Zahar's view threatens to make the confirmation of a new theory rather easy: for if none of the known facts which it entails played a role in its construction, then they will all confirm it irrespective of whether they also confirm existing theories. Yet scientists themselves regard genuine confirmations of, say, Relativity Theory as very difficult to come by indeed—which suggests that by genuine confirming evidence for a new theory they mean evidence which is not also explained by existing theories.

Lakatos uses the idea that background knowledge should be replaced by a background theory to simplify and improve Popper's definition of the severity of tests. The severity of a test resulting in $e$ on the hypothesis $h$ in the light of the 'touchstone theory' $t$ can be given by:

$$S(e, h, t) = p(e, h) - p(e, t).$$

As this formula makes clear, according to the third variant confirmation depends solely upon the logical relations between the evidence and the competing theories in question. Facts about the man who proposed the new theory, or about the heuristic route he followed, will be quite irrelevant.

In view of this, one may wonder whether the third view is really entitled to be called a variant of the historical approach at all. And it certainly is possible to use it as a theory of the selective confirmation provided by the evidence $e$ for the theories $T_1$ and $T_2$. Used in this way, the background theory $T_2$ to a given theory $T_1$ can be chosen quite arbitrarily.

---

1 Lakatos [1968], p. 382. Where both $h$ and $t$ predict $e$, $S(e, h, t)$ is 0. Where $h$ predicts $e$ but $t$ does not, $S(e, h, t)$ is $1 - r$ where $r$ is $p(e, t)$. Where $h$ predicts $e$ but $t$ predicts not-$e$, so that $p(e, t)$ is 0, $S(e, h, t)$ is 1. Thus Lakato's definition gives maximum severity to crucial experiments between $h$ and $t$, while Popper's places no premium on crucial experiments at all (see above, note 1 to page 7).

One small anomaly can be easily repaired. If $h$ predicts $e$ and $t$ predicts not-$e$, and if we let the evidence statement of the formula be not-$e$, it yields a severity of minus 1. So we should take the absolute value of $p(e, h) - p(e, t)$ to give us severity.
But our third view can also be used as a variant of the historical theory, provided we require that the 'background theory' for a new theory must be the best available theory actually present in the field. Then it will be a historical task to locate that theory in any given case. And if it is used in this way, our third variant, like the first, introduces a historical relativity into confirmation.

To see this, let us return to our three schematic situations. In situation (1), $T$ will be confirmed by $e_2, \ldots, e_n$ since it predicts them and (by assumption) no existing theory does the same. But in situation (2), evidence $e_2, \ldots, e_{n-1}$ will not confirm $T$ because all of it is also predicted by the background theory $T_{n-1}$. In situation (3), only $e_n$ confirms $T$—because it refutes its background theory. Thus in situation (3) $T$ will not be confirmed by evidence which might have confirmed it had the historical situation been different, and $T$ not been preceded by $T_{n-1}$. For example, according to this third variant, because Einstein had the misfortune to be preceded by Newton, his theory cannot be confirmed by all the evidence which it predicts, but which is also predicted by Newton’s theory.

Our schematic situations do pose one minor problem. Does $e_1$ confirm $T$ in situation (1)—and do $e_1, \ldots, e_n$ confirm $T$ in situation (2)? In other words, what are we to regard as the ‘background theory’ to the first testable theory in a given field? The natural answer, it seems to me, is to assume that in such a case the background theory is the empty or tautologous theory. If we do so, then the first testable theory in a given field will be confirmed by all the phenomena which it explains (whether or not they were known in advance, and whether or not they played a heuristic role in the construction of the theory). The Bohr–Balmer case may be a case in point: if Bohr’s theory was the first explanation of Balmer’s empirical formulas, then it can be confirmed by them.\(^1\) We see, then, that the first theorist in a given field has a marked advantage as far as evidential support is concerned—while the better supported are existing theories, the more difficult it is to improve upon them. This is a point to which I shall return.

To sum up our discussion of the three variants of the historical theory. The first variant takes ‘background knowledge’ to include everything ‘known to science’ prior to the proposal of the new theory—it has the counterintuitive result that a new theory cannot be confirmed by known facts which refute its predecessor. The second variant takes ‘background

---

\(^1\) In this way we arrive at the same result as Lakatos (see above, p. 11) but for a very different reason. Incidentally, Balmer’s spectroscopic investigations seem to refute the Popperian historical conjecture I mentioned earlier (see above, p. 10). For they were not undertaken as tests of any prior theory. Balmer did, of course, make use of all sorts of theories (for example, those involved in the construction of the spectroscope)—but to make use of a theory and to test it are not the same.
knowledge’ to include only known things which played a heuristic role in the construction of the new theory—it has the counterintuitive result that how well a theory is confirmed might depend upon who we take to have proposed it. The third variant takes ‘background knowledge’ to include only the best existing competing theory—and for my money it is the best version of the historical approach to confirmation.

Whichever variant we prefer, it will have the result that a theory need not be empirically confirmed by evidence which shows that one of its predictions is true. In case anybody still cannot stomach this consequence, I shall conclude by trying to make it a little more palatable.

6 REQUIREMENTS FOR THE GROWTH OF KNOWLEDGE

To do this, we need only remember that empirical confirmation is just one of the requirements for a good theory. Popper, for example, lists three requirements that a new theory must satisfy if it is to represent progress in science. The first requirement is that the theory be simple or unified. The second is that it be independently testable or not ad hoc. And the third is that it be (independently) confirmed—it ‘should pass some new and severe tests’.1

Now we have seen that the notions of an independent test and an independent confirmation can be interpreted in (at least) three different ways—and it is not clear from Popper’s formulations which he intends.8

Moreover, Popper’s third requirement has occasioned some controversy. He actually requires, not merely that the new theory be independently confirmed, but that it pass the first independent tests to which it is subjected. I agree with Lakatos that this was careless of him.3 Furthermore, Popper claims, rather mysteriously, that his third requirement is necessary to eliminate ad hoc theories:

‘... it is clear that the mere fact that the theory is ... independently testable cannot as such ensure that it is not ad hoc. This becomes clear if we consider that it is always possible, by a trivial stratagem, to make an ad hoc theory independently testable, if we do not also require that it should pass the independent tests in question: we merely have to connect it (conjunctively) ... with any testable but not yet

1 Popper [1963], pp. 240–8. If we construe ‘independently testable’ in accordance with the third variant, then Lakatos’s ‘acceptability,’ and ‘acceptability2’ (cf. his [1968], pp. 375–90) correspond to Popper’s second and third requirements respectively.

2 For example, Popper’s formulation of his second requirement is ambiguous between the first two variants of the historical theory: ‘For, secondly, we require that a new theory should be independently testable. That is to say, apart from explaining all the explicanda which the new theory was designed to explain, it must have new and testable consequences (preferably consequences of a new kind); it must lead to the prediction of phenomena which have not so far been observed’ ([1963], p. 241). Later Popper hints also at the third variant: see, for example, his [1963], pp. 246–7.

3 Cf. Lakatos [1968], p. 388.
tested fantastic *ad hoc* prediction which may occur to us (or to some science fiction writer).

Thus our third requirement, like the second, is needed in order to eliminate trivial and other *ad hoc* theories.¹

This is a bad argument for a good principle. First of all, the term *'ad hoc'* is used in two very different senses (at least—testable predictions are also described as *'ad hoc'*). First, a theory is *ad hoc* if it is not independently testable. Second, an independently testable theory is *ad hoc* if it is not independently confirmed.²

But more important than the obscurity of the argument is the fact that Popper's third requirement cannot do the job Popper here gives it, nor is it necessary for it to do so. For suppose we do conjoin a theory *T* which is not independently testable with an arbitrary testable prediction *P*. And suppose that the prediction gets confirmed. Then the *'theory' T & P will satisfy Popper's third requirement. I think, however, that scientists would regard this as an unsatisfactory *'theory', and would not even bother to submit it to test. And they would do so because it does not satisfy the requirement that Popper places first in his list, the requirement of simplicity. Popper formulates this as follows:

The new theory should proceed from some *simple, new, and powerful unifying idea* about some connection or relation . . . between hitherto unconnected things . . . or facts . . . or 'theoretical entities' . . . This *requirement of simplicity* is a bit vague, and it seems difficult to formulate it very clearly . . . [It is] intuitively connected with the idea of a unified or coherent system or a theory that springs from *one* intuitive picture of the facts . . .³

Now by assumption the arbitrary conjunction *T & P* is not simple or unified or coherent, and will already be excluded by Popper's first requirement. We do not need his third requirement to do the job, nor can it really do so, as we have seen.

Are we then to dispense with Popper's third requirement, since the argument he gives for it is a bad one? I do not think so, for Popper gives another very simple argument in its favour. Scientists search for truth, and want a new theory to have more truth in it than its predecessor. And a new theory will not have more (empirically checkable) truth in it than its pre-

¹ Popper [1963], p. 244.
² Lakatos labels these two senses of the term *'ad hoc'*, and *'ad hoc\(_2\)'* respectively (in his [1968], p. 389, note 1). I shall not follow him in this; I see no point in calling a theory *ad hoc\(_2\)* when what you mean is that it has not been independently confirmed. Alas, having introduced the subscripts, Lakatos warms to the device and makes the terminological thicket even thicker. He introduces the term *'ad hoc\(_3\)'* to describe theories (or rather, theoretical developments) which are not simple or unified (cf. his [1970], p. 175). Zahar goes even further, and suggests, in effect, that we add a further temporal subscript to Lakatos's *'ad hoc\(_3\)'* (see his [1973], p. 101, note 1).
³ Popper [1963], p. 241 (I have amalgamated the text and a footnote to it).
decedor unless at least one of its independent predictions gets confirmed. Hence the requirement of independent confirmation.\footnote{For this second argument see Popper's [1963], pp. 245–6. In stating it I made two large assumptions. First, that the scientist's decisions about the success or failure of predictions are correct. Second, that the new theory does contain all the true testable consequences of the old. I have also avoided the very real difficulties in giving an adequate definition of Popper's notion of verisimilitude. As Tichý has shown, if the new theory and the old are both false, then on Popper's own definitions the one cannot have greater verisimilitude than the other—see Tichý [1974].}

Now I promised to use these three requirements to try to make the historical approach to confirmation a little more palatable. It is very easy to do this. We have three requirements for a good theory—and ideally, of course, a new theory will satisfy all three of them. But the requirements are exacting ones, and the more successful science has been in fulfilling them to date the more difficult it becomes to fulfil them in the future. The deeper and more comprehensive our present theories, the more difficult it is genuinely to unify them into a new theory which is not only independently testable, but which passes some of these tests.\footnote{Newton is supposed to have remarked: 'If I have seen further it is by standing on the shoulders of giants'. Usually this is seen as an expression of his modesty—sometimes of his false modesty. I suspect that it may have been a straightforwardly megalomaniac remark: that it is more difficult to see further than a giant than it is to see further than a dwarf, and that it was very clever of him to have done so.} I suggest, therefore, that failure to satisfy all these requirements should not be regarded as a condemnation of a theory. This is pretty trite. A theory, independently testable or not, which genuinely unifies previously disjoint facts or theories can be a theoretical triumph, because of the theoretical problems it solves and the theoretical suggestions it makes. An independently testable theory which never gets confirmed can still contribute by leading to new factual discoveries and posing new problems for the theoretician. Of course, the best theories will satisfy all three of the requirements.\footnote{I do not think Popper would disagree with all this. He says that an ad hoc theory which represents a genuine unification may be a very great achievement ([1963], p. 241). And he says that a theory which satisfies his first two requirements but not the third can be 'an important contribution to science' ([1963], p. 243; his example is the Bohr–Kramers–Slater theory).}

At any rate, I do not think that historical theorists can have it both ways. By allowing only independent tests (however construed) to confirm, they make confirmation more difficult. And it becomes more and more difficult the more successfully science has met their requirements in the past. So they should not be too hard on theories which, by their austere standards, are ad hoc.\footnote{Of course, there is always a suspicion that an ad hoc theory has been merely 'cooked up' in some trivial way to yield existing results. But an independently testable theory may also have been 'cooked up'—and so may an independently confirmed one (see above, p. 20). Philosophers of science have made little progress in explaining how a 'cooked up' theory}
scientific virtues—but like moralists they should not be too swift to condemn that which has only some of them.\footnote{1}

7 CONCLUSION

If I were asked to sum up, I would say this. The logical approach to confirmation renders it so easy that it ceases to be important. While the historical approach renders it so difficult that it should cease to be all-important.

University of Otago,
Dunedin, New Zealand

REFERENCES


GOODMAN, N. [1954]: Fact, Fiction and Forecast.

differs from a coherent or unified one (though not much sophistication is needed to recognise that the results of the philosopher’s tacking arguments are ‘cooked up’). But working scientists do have means of distinguishing them.

\footnote{1}{The fact that there is more than one desideratum for a good theory (and hence more than one associated methodological rule) has some interesting repercussions. For in a given case the desiderata or rules may give conflicting results: \( T_2 \) may be more coherent than \( T_1 \), while \( T_1 \) is better supported by evidence than \( T_2 \). In such a situation two scientists might make their choices in different ways. And (short of attempting to weight his desiderata) the methodologist will have to say that both of them are acting rationally. Thomas Kuhn has been making the point for years—it is actually all that his famous ‘incommensurability thesis’ really amounts to (see my [1971], p. 294).

I am also tempted to try to bring Feyerabend back from beyond the irrationalist pale. He claims to be an epistemological anarchist. He says that, for any methodological rule yet proposed, he will find a scientist who violated it—and a good thing too. And he concludes that the only sensible rule is: there are no rules, anything goes.

But if there are multiple desiderata (multiple rules), and they give conflicting results, an individual may quite rationally set one aside in favour of another. Inspection of Feyerabend’s case-study of Galileo (in his [1970]) will, I think, reveal that this is exactly what happened: Galileo was no ‘anarchist’ but his rules gave him conflicting results and he had to back some and violate the others temporarily. Thus Feyerabend’s premise (each rule has a fruitful historic violation) does not, even if correct, establish his conclusion (there should be no rules). Similarly one cannot establish moral anarchism by showing that, for any moral rule, there are situations where it ought to be violated—what this would show is only that no single rule is so sacrosanct that it overrides all others in all situations.

Feyerabend’s ‘epistemological anarchism’ has taken him so far to the left of the other Popperians that he is in danger of falling off their platform altogether. The above will rescue him from this sad fate. He will, no doubt, be furious at the rescue act, and may make even more frantic efforts to fall off. Which is why I try to rescue him, for his cavortings are always amusing and sometimes instructive.}
HEMPEL, C. G. [1965]: Aspects of Scientific Explanation and other essays in the Philosophy of Science.
KEYNES, J. M. [1921]: A Treatise on Probability.
MEDAWAR, P. [1967]: The Art of the Soluble.
MILL, J. S. [1843]: A System of Logic.
WHEWELL, W. [1840]: Philosophy of the Inductive Sciences.