PREDICTIVISM

Richard Swinburne

Oriel College, Oxford

Abstract. Normally, whether evidence of observation e confirms hypothesis h relative to background evidence k is a matter of the logical relations between e, h, and k, and independent of whether h was formulated before or after the discovery of e. This is supported by intuitions about various examples from history of science, and fits in with a Bayesian account of confirmation. There are however abnormal cases where the background evidence has a special character, when it does become relevant when h was formulated and k records that.

"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 ..." wrote Alan Musgrave in 1974. He called "this cornerstone of logical empiricism the ... logical ... approach to confirmation", contrasting it with the historical approach, which holds that the order or circumstances of discovery or formulation of e or h are relevant to confirmation. Clearly we must add k, background knowledge to the equation, and then the logical approach will hold that confirmation is a matter solely of the logical relations between h, e, and k, and independent of the order or circumstances of the discovery or formulation of the different elements. But even with this obvious addition, I suggest that Musgrave has misstated the issue somewhat. Confirmation is a measure of the support given by evidence to hypothesis, and the only thing that can function as evidence is something known. The order or whatever of discovery is clearly irrelevant to confirmation unless it is known, and then it itself will be an element of evidence. Confirmation is therefore inevitably (in a wide sense) a logical matter. The dispute is rather between those who consider that when we have some evidence e_1 relevant to the confirmation of a hypothesis h, the addition to e_1 of further evidence e_2 to the effect that e_1 was in some sense novel evidence, discovered in some sense subsequent to the

formulation of h increases the value of the confirmation, and those who think that e_2 is irrelevant. I shall call the two views the timeless view and the historical (or predictivist) view. Since I hold that these matters can be formalised by the probability calculus and in particular by Bayes's theorem, I shall understand 'confirmation' in its terms. Thus restricted, it can still be understood either as 'probability' ($P(h/e\&k)$) or as increase of probability in absolute terms ($P(h/e\&k) - P(h/k)$) or proportional increase ($P(h/e\&k)/P(h/k)$). Clearly the results for our controversy will be the same, in whichever of these ways we understand 'confirmation'. I shall therefore understand it simply as 'probability'. The considerations which I shall adduce will, I think, normally also hold on understandings of confirmation which cannot be axiomatised in terms of the calculus, which I shall call non-Bayesian understandings; though I shall not argue for that directly. The issue therefore boils down to – Is $P(h/e_1\&e_2\&k)$ ever different from $P(h/e_1\&k)$, where e_2 is evidence of the novelty of e_1.

There are three different understandings of the novelty of e_1 and so of the priority to it of h. e_2 might report the temporal priority of the formulation of h to the discovery of e_1; e_1 is then 'temporally novel'. Then there is the psychological priority of the formulation of h; e_1 is not taken into account in the formulation of h and so e_1 is 'use novel'. And finally there is the epistemic priority of the formulation of h; people at the time had no good reason provided by current theories for believing that e_1 would occur before h was put forward. This I shall call the 'relative novelty' of $e_1 - e_1$ is relatively novel if, while it is quite probable given h and k, it is very improbable given the actual theories in vogue when h was formulated. This 'relative novelty' is a historical feature of the circumstances of the discovery of e_1 (in what climate h was formulated) – to be distinguished sharply from the low prior probability of $e - P(e/k)$, or its low probability if h is false – $P(e/\sim h\&k)$. The latter two values arise from the probability of e on all the various hypotheses which might hold, weighted by their prior probability – whether or not they were recognised at the time of the formulation of h. Of these three kinds of novelty, most writers discount temporal novelty as the relevant kind, preferring either use novelty (e.g. Worrall [1989]) or relative novelty (Musgrave [1974]). I shall call any such evidence of the novelty of e_1 novelty evidence. It is a species of historical evidence, in the sense of evidence about which hypotheses were formulated by whom and when and under what circumstances. I shall argue that for normal k novelty evidence is irrelevant to the confirmation of h, but that for certain k such e_2 is relevant.

By normal k, I mean that k describes the experimental conditions in which e_1 occurs (not the conditions in which it was observed) or background evidence about other scientific theories which are relevant to whether e_1 confirms h. k is not to include any historical evidence about who formulates h or other hypotheses under what circumstances and what their success rate is.

Here is a trivial example in which k is normal and in my view the timeless theory is correct. Let h be 'all metals expand when heated,' k be '1000 pieces of

metal were heated', e_1 be 'those 1000 pieces of metal expanded', e_2 be some novelty evidence such as that h was put forward before e^1 was known. k and h (if true) would explain e_1 – the metals expanded because they were heated, and because all metals expand when heated. (I assume that the latter is a lawlike connection). The timeless theory then claims that $P(h/e_1\&k) = P(h/e_1\&e_2\&k)$. I suggest that the simplicity of the hypothesis and its ability to explain (given k) a large amount of evidence is what makes it likely to be true, quite independently of when it was put forward. It is also quite independent of what led anyone to formulate h, and of whether e was very probable or very improbable given the theories in vogue at the time of the formulation of h. Some very crazy astrologically based theories might have been in vogue then which (together with k) predicted and explained e_1. Yet that would not make e_1 any less good evidence for h.

In the above example h is a universal hypothesis. The timeless theory also works for predictions – given normal k. Observational evidence is typically evidence for predictions by being evidence confirmatory of universal (or statistical) hypotheses of which the prediction is a deductive (or inductive) consequence. Let e_1 and e_2 be as before, k be '1001 pieces of metal were heated', and h be 'the 1001th piece of metal expanded. e_1 with k is evidence that 'all metals expand when heated, and so that the 1001th piece of metal will expand when heated. h derives its probability from being a consequence of a simple hypothesis able to explain a large amount of evidence, independently of whether or not e was novel in any sense. $P(h/e_1\&k) = P(h/e_1\&e_2\&k)$.

I report my intuitions on this simple example. They are also my intuitions on more sophisticated real-life examples, where my condition on background knowledge holds. They are, for example, my intuitions with respect to Mendeleev's theory, recently discussed by Maher [1993] on the historical side, and by Howson and Franklin [1991] on the timeless side. Mendeleev's theory (h) entailed and (if true) explained the existence and many of the properties of scandium, gallium, and germanium (e_1). Mendeleev's was not just any old theory which had this consequence; it was not just e_1 plus some unrelated f. It was an integrated theory of groups of related elements having analogous properties recurring as atomic weight increased, from which the existence of the sixty or so elements already known (k) followed. In virtue of being a more integrated and so simple theory than any other theory from which that followed and by which it could be explained, it was already more likely to be true than any other theory. The further information that other results followed from it and could be explained by it was therefore plausibly further evidence for it independently of when and how they were discovered. Howson and Franklin compare the relation of Mendeleev's theory to chemical elements to the relation of the eightfold way to elementary particles; and the prediction of the three elements by the former to the prediction of the Ω-particle by the latter. They ([1991] pp. 579–80) cite a passage from Yuval Néeman, one of the inventors of the eightfold way in which he also

makes the comparison and comments that "the importance attached to a successful prediction is associated with human psychology rather than with scientific methodology. It would not have detracted at all from the effectiveness of the eightfold way if the Ω – had been discovered before the theory was proposed."

That theories can acquire a very high degree of support simply in virtue of their ability to explain evidence already available is illustrated by the situation of Newton's theory of motion at the end of the seventeenth century. It was judged by very many – and surely correctly – to be very probable when it was first put forward. Yet it made no new immediately testable predictions, other than the predictions which were already made by laws which were already known and which it explained (e.g. Kepler's laws of planetary motion and Galileo's law of fall). Its high probability arose solely from its being a very simple higher-level theory from which those diverse laws are deducible. My intuitions tell me that it would have been no more likely to be true, if it had been put forward before Kepler's laws were discovered and had been used to predict them.

So much for my intuitions. But my intuitions clash with those of the predictivist. So I need to show that my intuitions fit into a wider theory of confirmation for which other reasons can be adduced, and I need to explain why the predictivist has the inclination to give a wrong account of cases such as I have cited. My intuitions fit into the whole Bayesian picture, in favour of which of course there are other good reasons. Consider my first simple example in which h = 'all metals expand when heated', k = '1000 pieces of metal were heated', e_1 = 'those 1000 pieces of metal expanded', and e_2 = 'h was put forward before e_1 was known." On Bayes's theorem $P(h/e_1 \& k) = \dfrac{P(e_1 / h \& k)}{P(e_1 / k)} P(h / k)$ and

$P(h / e_1 \& e_2 \& k) = \dfrac{P(e_1 \& e_2 / h \& k)}{P(e_1 \& e_2 / k)} P(h / k)$. Then $P(h/e_1 \& e_2 \& k)$ will only be

greater than $P(h/e_1 \& k)$, as the predictivist claims it is if the addition of e_2 to the evidence lowers $P(e_1/k)$ (and so $P(e_1/{\sim}h_1 \& k)$ by a greater proportion than it lowers $P(e_1/h \& k)$. What this would mean is that on the mere information that 1000 pieces of metal were heated, it would be more likely that the hypothesis that all metals expand when heated would have been proposed before it was known that 1000 pieces of metal expanded if the hypothesis was true, than if it was false. That seems to me, and I hope also to the average predictivist, massively implausible. The same applies if we take e_2 as 'e_1 was used in formulating h_1.' Then a Bayesian predictivist is committed to: on the mere information that k it would be more likely than when e_1 was discovered, e_1 would be taken into account in formulating h if h were true than if it were not. And if e_2 is 'the theories in vogue at the time of the formulation of h were such that e_1 is improbable given them', the Bayesian predictivist is committed to: it would be more likely than when e_1 was discovered, the theories then in vogue would predict e_1 if h were true than if it were not. All of this is again massively implausible. Hence the predictivist can only save his thesis

by abandoning Bayes's theorem; and there are, I suggest, good reasons for not doing so.

But what, if predictivism is false, is the source of the temptation to espouse it? I think that there are two sources. First, there is the consideration that any collection of evidence can always accommodate some theory i. e. for any e_1, and k one can always devise a theory h, such that $P(e/h\&k) = 1$, or is high. (Indeed, one can always devise an infinite number of such theories.) That has seemed to suggest, totally erroneously, to some that there are no objective criteria for when a theory so constructed is supported by evidence. Whereas, the contrast is made, once we have a hypothesis which makes a prediction, we can look to see whether the prediction comes off and whether it does or not is a clear objective matter. But in fact there are normally very clear objective criteria for when a theory is supported by a collection of evidence. In the trivial metal example which I used earlier, equally accommodating to the evidence that 1000 pieces of metal had been heated and expanded, are h "all metals expand when heated," h^1 "metals 1–1000 expand when heated, and other metals do not" and, supposing all the metals observed so far were observed by physicists in Britain, h^{11} "All and only metals observed by physicists in Britain expand when heated." Quite obviously, h^1 and h^{11} are not supported by evidence, whereas h is. The obvious reason for this (though some philosophers of science seem to make a lot of heavy weather about recognising this) is that h is (relative to the evidence and logically possible rivals) a simple hypothesis, whereas h^1 and h^{11} are not simple hypotheses.

Simplicity is a matter of a hypothesis being mathematically simple, postulating few entities and kinds of entity, few properties and kinds of property (fitting with background knowledge) and thus being integrated and not ad hoc. Of course it is a difficult philosophical task (not yet achieved, in my view) to spell out in detail just what this amounts to. But that casts no doubt on the fact that this criterion is at work, as so obviously it is; and that it is a criterion of truth – simpler theories, as such, are more probably true. If we didn't think that, we would think that on the evidence cited, h^1 and h^{11} would be likely to be true as h, and so their predictions were equally likely to be true – which obviously we don't think. Theories which "fit better" with background knowledge, knowledge of the laws operative in neighbouring fields, are also more likely to be true. But this is so ultimately because of the same criterion of simplicity – theories fitting better with background knowledge constitute simpler theories for the total area (the new area and its neighbouring fields) than do theories which do not fit well. Simpler theories, as such, have greater prior probability. So too, as everyone recognises, do theories with less content – i.e. theories which make less and less precise claims about the world – if h_1 entails h_2, then for any e and k $P(h_1/e\&k) \leq P(h_2/ehk)$, and normally '$\leq$' is to be read as '<'. The more you say, the more details you commit yourself to, the more likely you are to be mistaken.

So there are in general clear objective criteria for when a theory which accommodates the evidence is supported by it. Not always however. Our criteria

are not that sharp, and some scientists understand simplicity a little differently from others. And in such cases rival hypotheses are equally well supported by evidence (or, even if one is better supported, it is not evident that it is). But this fuzzy range of unclarity need cast no doubt on the main point that not all evidence can be accommodated by a hypothesis sufficiently simple for the evidence to support it. There are indeed objective texts for which of two theories is supported by prior evidence, which usually give clear results. And particular predictions may give no sharper results – they may discriminate between many theories, but they cannot discriminate between all theories (only an infinite number of predictions could do that).

These points about more precise claims having as such lower prior probability and simpler theories having as such higher prior probability explain predictivist intuitions about another kind of example which they sometimes adduce, where subsequent observation may either fill in the details of an existing theory or be predicted by an existing theory. Here, they say, prediction is better evidence of truth than accommodation. We have a theory h_1 with a variable parameter. e_1 is observed. h_1 will hold if and only if that parameter has a certain value. Giving that parameter that value turns h_1 into a more precise hypothesis h_2. Contrast this with another equally precise hypothesis h_3 formulated before e_1 was observed, which successfully predicts e_1. Then, says the predictivist, $P(h_3/e_1\&k) > P(h_2/e_1\&k)$. But – assuming $P(e_1/h_3\&k) = P(e_1/h_2\&k)$ – that would only be the case, if in advance of e_1 being observed, says the Bayesian, $P(h_3/k > P(h_2/k)$. That would be the case if $P(h_3/k) = P(h_1/k)$, because the more detailed filling out of h_1 will yield a theory with lower prior probability. And that is what the predictivist has assumed in his example – that really there is nothing to choose between h_1 and h_3 before e_1 is observed (e.g. because although h_3 is more precise than h_1, it is also simpler). But in that case certainly h_3 will be better confirmed by e_1 than h_2 for reasons which have nothing to do with the novelty of e. But if, to start with, $P(h_1/k) > P(h_3/k)$, there is no reason to go along with the predictivist's intuitions about the example.

The second source of predictivism, as I see it, is this. A hypothesis h which entails (for some circumstances k) the occurrence of some event e, has its probability raised by e, the less likely e is to occur if h were not true – the lower $P(e/\sim h\&k)$ and so the lower $P(e/k)$. If we have already formulated h_1 we know which e to look for which will have this characteristic of $P(e/h\&k) = 1$ and $P(e_1/\sim h\&k)$ very low. We can bring about k and see whether $\sim e$ or e occur, and that will provide a "severe test" of the theory. If we formulate h after accumulating evidence, we may or may not have among that evidence an e with that characteristic – but we are much more likely to have it if we are actually looking for it. Hence producing hypotheses and then testing them may indeed be a better way of getting evidence which (if they are true) supports them strongly, than trying to fit them to evidence we already have. But that has no tendency to cast doubt on the fact that for given evidence e, $P(h/e\&k)$ has a value independent of when e was discovered. There is, it is true, always the temptation for the

accommodator to include among his evidence only that which the hypothesis leads him to expect. But the probability which ought to guide action is that relative to total available evidence – and all relevant evidence should be taken into account, and that will include evidence about any methods used to obtain other evidence (e.g. optional stopping such as ceasing to toss your coin when the proportion of heads to tails reaches some preferred value). This second source of predictivism is the criterion of the 'severe test' which Mayo [1991] among others, has correctly seen as lying behind the demand for novel evidence; but, as she has argued, novel evidence in no sense is either necessity or sufficient for severity of tests.

My claim that novelty evidence e_2 is irrelevant to confirmation, and my explanation of why some writers have thought otherwise, is only supposed to apply in the normal sorts of case. These are the cases where the background evidence k is the normal evidence about the circumstances of the occurrence of e_1 (not the circumstances in which it was observed) and evidence about the worth of other scientific theories relevant to whether e_1 confirms h. But my claim does not hold for any k and it would be very odd to suppose that it did – since for any evidence and any hypothesis, there is always some background evidence which makes the former relevant to the latter. In particular my claim does not hold in many cases where k reports evidence of a historical kind, the force of which (together with e_1 and e_2) is to indicate that someone has access to evidence relevant to h which is not publicly available.

Here is an example where the k is evidence of this historical kind. Let h be Grand Unified Field Theory, and e_1 be some observed consequence thereof. Let k be that h was formulated by Hawks who always puts forward his theories after assembling many pieces of observational evidence which he does not reveal to the public, and which – so long as, subsequent to being formulated, they make one true new prediction – are always subsequently confirmed and never falsified. Then of course the evidence that a consequence of the theory was observed subsequent to its formulation, increases its confirmation. $P(h/e_1 \& k) < P(h/e_1 \& e_2 \& k)$. Or - another example – let h be a theory about the order of cards in a pack of cards, e_1 be that the pack was shuffled by the well-known card – sharper Jones, k that h was propounded by a rival card – sharper Smith who bet on h. Since Smith is more likely to know the order of the cards after they have been shuffled than before – for if the bet is placed before the shuffle Jones will try to arrange them in an order other than h – $P(h/e_1 \& e_2 \& k) > P(h/e_1 \& k)$. In both of these examples, the historical evidence is evidence which (together with e_1 and e_2) shows that someone knows more about h, than merely e_1; they have evidence not accessible to the public and are for that reason to be trusted (or, for some k, not to be trusted).

In the light of all these considerations, let us turn to the example put forward by Maher in defence of predictivism:

*We imagine an experiment in which a coin is tossed 99 times, and a subject
records whether the coin landed heads or tails on each toss. The coin seems
normal, and the sequence of tosses appears random. The subject is now asked to
state the outcome of the first 100 tosses of the coin. The subject responds by
reading back the outcome of the first 99 tosses, and adds that the 100th toss will
be heads. Assuming that no mistakes have been made in recording the observed
tosses, the probability that the subject is right about these 100 tosses is equal to
the probability that the last toss will be heads. Everyone seems to agree that
they would give this a probability of about 1/2.*

*Now we modify the situation slightly. Here a subject is asked to predict the
results of 100 tosses of the coin. The subject responds with an apparently
random sequence of heads and tails. The coin is tossed 99 times, and these
tosses are exactly as the subject predicted. The coin is now to be tossed for the
100th time, and the subject has predicted that this toss will land heads. At this
point, the probability that the subject is right about all 100 tosses is again equal
to the probability that the 100th toss will land heads. But in this case, everyone
seems to agree that they would give it a probability close to 1.*

*The difference between the two situations is that in the first the subject has
accommodated the data about the first 99 tosses, while in the second that data
has been predicted. Clearly the reason for our different attitude in the two
situations is that the successful prediction is strong evidence that the subject has
a reliable method of predicting coin tosses, while the successful accommodation
provides no reason to think that the subject has a reliable method of predicting
coin tosses.* (Maher [1993] p. 330)

Let e_1 be the outcomes of the first 99 tosses, h be e_1 plus the proposition that
heads will occur on the 100th toss, e_2 be that h was formulated before e_1 was
observed, and k be a description of the set-up "where the coin seems normal and
the sequence of tosses appears random." k will also have to include the
information that h was the only (or almost the only) hypothesis formulated, for as
Howson and Franklin ([1991] p. 577) point out if all possible hypotheses have
been formulated, the example won't work. h would be no more likely to be true
than the hypothesis consisting of e_1 plus the proposition that tails will occur on the
100th toss. The fact that someone guessed the lottery numbers correctly is no
reason for supposing that he will guess the numbers correctly next time, when on
the successful occasion all possible numbers had been guessed by someone or
other.

However, given k as above, claims Maher, "everyone seems to agree that"
$P(h/e_1\&e_2\&k)$ is close to 1. Everyone is surely correct on this. Yet, Maher also
claims, "everyone seems to agree" that $P(h/e_1\& \sim e_2\&k)$ is about ½. Hence it will
follow that $P(h/e_1\&e_2\&k) > P(h/e_1\&k)$, and so historical evidence increases
confirmation. But if "everyone agrees" to the latter, it is just possible that they
have been conned. The apparently random may not be really random. It may be
that there is a pattern of regularity in the first 99 tosses, which the subject who put
forward h has alone spotted. Then $P(h/e_1\&k)$ will also be close to 1 (even though

most of us are too stupid to realise that), and the historical information e_2 is irrelevant.

But suppose there is no pattern in the tosses. In that case what "everyone agrees" about the probabilities is correct. So we ask why is $P(h/e_1\&e_2\&k)$ close to 1. The answer is that k includes historical information that h was the only hypothesis put forward. That together with e_1 and e_2 – the fact that his predictions were so accurate – is very strong evidence that the hypothesiser has access to information about bias in the set-up that we don't (either via some publicly observable evidence or some private intuitions – maybe he has powers of telekinesis). This is for the reason, given by Howson [1988] pp. 383–4 that $(e_1\&e_2\&k)$ would be very improbable if the hypothesiser did not have this information. That is, we trust the prediction because of who made it, not because of when it was made. That that is the correct account of what is going on here can be seen by the fact that suppose we add to k irrefutable evidence that the hypothesiser had no private information, then we must conclude that his correct prediction of the first 99 tosses was a mere lucky guess and provides no reason for supposing that he will be right next time. I conclude that unless k also includes historical evidence, the mere evidence of novelty provided by e_2 does not affect the probability of a hypothesis. In this case alone where the historical evidence shows private information is the "method" by which the hypothesis is generated of any importance for its probability. Maher ([1993] p. 335) claims that "the introduction of the concept of a method is the main conceptual innovation in my account of the value of prediction". In general the method by which the hypothesis was generated is irrelevant to its probability on evidence. Whether or not Mendeleev's theory was generated "by the method of looking for patterns in the elements," its probability depends on whether it <u>does</u> postulate patterns, not how it was arrived at. Kekule's theory of the benzene ring is neither more or less probable on its evidence because it was suggested to Kekule in a dream. Only if the evidence suggests that someone has private information does it become important whether the hypothesis was generated after consideration of that information. Then given that (and only given that) we have reason to believe that the hypothesiser is a tolerable mathematical scientist and keen to formulate true theory, is the fact that it was produced by whatever method gave it success so far of relevance. Evidence of how a theory was produced is indirect evidence of whether on the non-public evidence it was likely to be true. But if we have the evidence for ourselves, we can ignore all that.

Address:
   Richard Swinburne
   Oriel College
   Oxford OX1 4EW UK
E-mail: richard.swinburne@oriel.ox.ac.uk
Phone: 01865-276555
Fax: 01865-791823

References

Howson, C (1989) "Accommodation, Prediction, and Bayesian Confirmation Theory". *Philosophy of Science Association Proceedings*, 1988 vol 2, 381–92.

Howson, C. and Franklin, A. (1991) "Maher, Mendeleev, and Bayesianism". *Philosophy of Science*, 58, 574–85.

Maher, P. (1993) "Discussion: Howson and Franklin on Prediction" *Philosophy of Science*, 60, 329–40.

Mayo, D. G. (1991) "Novel Evidence and Severe Tests". *Philosophy of Science*, 58, 523–552.

Musgrave, A. (1974) "Logical versus Historical Theories of Confirmation". *British Journal for the Philosophy of Science*, 25, 1–23.

Worrall, J. (1989) "Fresnel, Poisson, and The White Spot: The Role of Successful Predictions in the Acceptance of Scientific Theories". In *The Uses of Experiment*. D. Gooding, T. Pinch, and S. Schaffer, eds. Cambridge: Cambridge University Press.