Part IX

INTERPRETING SCIENTIFIC INFERENCE

https://doi.org/10.1086/psaprocbienmeetp.1988.1.192994 Published online by Cambridge University Press

Prediction, Accommodation, and the Logic of Discovery [1]

Patrick Maher

University of Michigan

1. Arguments against the predictivist thesis

It is widely believed that if a piece of evidence for a theory was known at the time the theory was proposed, then it does not confirm the theory as strongly as it would if the evidence had been discovered after the theory was proposed. I shall call this view the *predictivist thesis*. Those who have endorsed it include Leibniz (1678), Huygens (1690, preface), Whewell (1847 vol. 2, p. 64f.), Peirce (1883), Duhem (1914, ch. II, §5), Popper (1965, p. 241f.), Lakatos (1970, p. 123), and Kuhn (1977, p. 322). On the other hand, the thesis has been rejected by a number of philosophers, including Mill (1872 bk. III, ch. 14, §6), Keynes (1921, p. 305), Rosenkrantz (1977, p. 169f.), Horwich (1982, pp. 108–117) and Schlesinger (1987). Others, while not rejecting the predictivist thesis, nevertheless regard the justification of the thesis as problematic; these include Hempel (1966, p. 38) and Gardner (1982).

The view that the predictivist thesis is problematic stems in large part from a perception that this thesis is incompatible with a Bayesian philosophy of science. However, the arguments which are supposed to demonstrate this incompatibility are fallacious, as the following brief discussion will show.

One simple argument, stated by Hempel, Lakatos, Rosenkrantz and Gardner, begins by asserting that the probability of hypothesis H, given evidence E, is determined by H and E; and from this it concludes that the probability of hypothesis H is independent of whether E was known at the time H was proposed. But while the premise of this argument is true, the conclusion does not follow; for there is nothing to prevent us taking the total evidence for H to include not only the evidence E, but also information about whether or not E was predicted—as Simon pointed out in (1955). Letting O be the proposition that evidence E had been observed at the time H was proposed, a Bayesian formulation of the predictivist thesis would be that[2]

$$P(H|E\bar{O}) > P(H|EO).$$

Nothing in the above argument shows that this inequality cannot be true.

Horwich offers an insightful account of what the predictivist thesis asserts, and also two arguments which are intended to show that the thesis is false. One of these (the sec-

PSA 1988, Volume 1, pp. 273–285

https://doi.org/10.1086/psaprocbienmeetp.1988.1.192994 Published online by Cambridge University Press

ond in his presentation) shows correctly that a certain argument for the predictivist thesis is unsound; but of course, this does not prove that the thesis is false. So I shall concentrate here on his other argument which, if correct, would indeed demonstrate that the predictivist thesis is false. Horwich uses the notation

$$F = \text{our theory fits the data } D$$
$$H_T = \text{our theory is true}$$
$$H_R = \text{we required a theory that would fit the data } D.$$

Here H_R is taken to entail F. In these terms, the predictivist thesis can be stated as

$$P(H_T|F\bar{H}_R) > P(H_T|H_R).$$

Horwich claims that the probabilities on either side of this inequality are in fact equal, and thus that the predictivist thesis is false. In support of this, he asserts that "what is relevant to the probability of H_T, in the information H_R, is precisely that our theory fits the data D." But this assertion is merely a restatement of the anti-predictivist position, and so as an argument against the predictivist thesis it simply begs the question.[3]

Some predictivists, such as Popper and Lakatos, have claimed that evidence which was known at the time a hypothesis was proposed provides *no* support for the hypothesis. This claim is not entailed by the predictivist thesis, and I shall not be discussing it here, except to say that some arguments against the predictivist thesis rest on a failure to distinguish the two propositions. For example, Schlesinger (1987, p. 37) asserts that scientific practice does not accord with the predictivist thesis, and offers as support of this such facts as that Galileo's experiments on free fall were taken to confirm Newton's theory of gravitation. But this fact, and the others Schlesinger cites, are perfectly compatible with the predictivist thesis; what they tend to show is only that evidence known at the time a hypothesis was proposed can provide support for that hypothesis.

In fact, scientific practice accords well with the predictivist thesis, as I shall now illustrate with a discussion of the scientific reaction to Mendeleyev's periodic table of the elements.

2. A historical example: Mendeleyev's periodic table

By the middle of the 19th century more than 60 chemical elements were known, with new ones continuing to be discovered. For each of these elements, chemists attempted to determine its atomic weight, density, specific heat, and other properties. The result was a collection of facts that lacked rational order. Mendeleyev noticed that if the elements were arranged by their atomic weights, then valences and other properties tended to recur periodically. However, there were gaps in the pattern, and in a paper of 1871 Mendeleyev asserted that these corresponded to elements which existed but had not yet been discovered. He named three of these elements eka-aluminum, eka-boron, and eka-silicon, and gave detailed descriptions of their properties. The reaction of the scientific world was skeptical. But then in 1874 Lecoq de Boisbaudran found an element which corresponded to Mendeleyev's description of eka-aluminum, and which he called gallium. This was regarded as a remarkable event; it was the first time in history that a person had correctly foreseen the existence and properties of an undiscovered element. Confidence that Mendeleyev's other predictions would be confirmed increased markedly. Four years later, Nilson discovered an element which corresponded to Mendeleyev's description of eka-boron, and which he named scandium. Now chemists were expecting to find Mendeleyev's third element, though the Royal Society did not wait for that discovery, awarding Mendeleyev its Davy Medal in 1882. Mendeleyev's eka-silicon was discovered by Winkler in 1886, and named germanium.[4]

https://doi.org/10.1086/psaprocbienmeetp.1988.1.192994 Published online by Cambridge University Press

If scientists accord no special confirmatory value to predictions, then it is quite inexplicable why their confidence in Mendeleyev's predictions should have increased substantially after one or two of those predictions had been verified. There were 62 elements in Mendeleyev's table of 1871, so we could say that Mendeleyev's prediction of eka-silicon was initially made on the basis of evidence concerning 62 elements.[5] After the discovery of gallium, the prediction concerning eka-silicon was now backed by evidence concerning 63 elements; and after the discovery of scandium, it was backed by evidence concerning 64 elements. But the difference between these bodies of evidence is much too small by itself to account for the dramatically altered attitude to Mendeleyev's prediction of eka-silicon. The only plausible explanation is that scientists were impressed by the fact that these latter pieces of evidence were verified predictions rather than accommodated evidence.

3. Predicting coin tosses

We have now seen that the predictivist thesis has not been proved inconsistent with Bayesian theory, and that it accords with the history of science; but we do not yet have an explanation of why the predictivist thesis might be true. My main aim in the remainder of this paper will be to provide such an explanation.[6] I begin with a consideration of the following artificial example, which I think is suggestive.

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 to be normal, and the sequence of tosses appears random. The subject is now asked to tell us the outcome of the first 100 tosses of the coin. The subject responds by reading back the outcome of the first 99 tosses, and then adding the prediction 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. I invite you to consider what probability you would give to this being true.

Now consider a slightly different case. 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 all 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. In this case, what probability would you give to the 100th toss being heads?

People I have talked to say that in the first case they would give a probability of around 1/2 to the 100th toss landing heads, while in the second case they judge the probability to be near one. Their rationale is that in the first case the subject is probably just making a random guess, while in the second case the successful prediction of 99 tosses is strong evidence that the subject is not merely guessing, but rather has some reliable method for predicting the tosses.

The difference between the two cases is a difference between prediction and accommodation of evidence. For let E be the hypothesis that the first 99 tosses are as they have been observed to be, and H the conjunction of E together with the prediction that the 100th toss will land heads. In both cases, our subjects have asserted H, and in both cases there is evidence E to support that assertion; the difference is that in the first case the evidence E was accommodated, while in the second case it was predicted. The widely shared intuition is that this difference makes a difference to the probability of H, because the prediction of E in the second case is evidence that our subject is an extraordinarily good predictor of coin tosses, while the accommodation of E in the first case is no evidence for the reliability of the subject's predictions.

https://doi.org/10.1086/psaprocbienmeetp.1988.1.192994 Published online by Cambridge University Press

This suggests that perhaps in the scientific case too, successful prediction might confirm that the method used to generate the hypothesis in question is a reliable one, and thereby provide additional confirmation for the hypothesis. This account of the special value of prediction also has some historical support, as can be seen by returning again to our example of Mendeleyev's periodic table.

Mendeleyev's method was to look for patterns in the properties of the elements. He was by no means the first to do this, but the earlier efforts had not been very successful, and there was considerable skepticism about the value of this approach. Berzelius commented that apparent numerical relationships could frequently be found between elements, but that subsequent revisions of atomic weights would alter these relationships, and therefore it was not safe to make speculative assumptions.[7] In 1866, Newlands presented a paper to the Chemical Society which organized the elements according to a "Law of Octaves", and was asked facetiously if he had ever thought of classifying the elements in alphabetical order. His paper was returned by the editor of the *Journal of the Chemical Society* as unsuitable for publication (Ihde 1964, pp. 240–242). The method of looking for patterns in the elements was evidently thought to have no better than random chance of leading to correct conclusions. But the success of Mendeleyev's predictions led the method to be taken seriously. When a similar situation arose in particle physics in the next century, the method of looking for patterns was judged likely to be a good method for trying to make sense of the data.[8]

In the next two sections I will show how this intuitive explanation of the value of prediction can be formalized within Bayesian confirmation theory.

4. Notation and Assumptions

I shall begin by reformulating the predictivist thesis in a way that brings out the role of the method by which the hypothesis was generated. To that end, let M be the hypothesis generation method operative on the occasion at issue, and let M_H be the proposition that M generated hypothesis H at time t. Let O be the proposition that the truth value of E had been input to M at time t. (This interpretation of O is weaker than the one we used in section 1, since on the present interpretation O does not imply that E is true.) Then the predictivist thesis can be stated as

$$P(H|M_H E\bar{O}) > P(H|M_H EO). \qquad (1)$$

The reliability of method M can be taken to be the probability that a hypothesis is true, given that it has been generated by M. Since we want to allow that we may be ignorant of the true reliability of a method, the probability referred to here will need to be an objective probability, not a subjective probability.

Now consider the special case in which method M is known to be either completely reliable, or else random. That is, the probability of a hypothesis being true, given that it was generated by M, is either 1, or else is the probability of a randomly chosen hypothesis being true.[9] What I will do here is give a Bayesian analysis of this special case. While this special case is not very realistic, the analysis which I shall give is capable of being generalized to more realistic cases. I concentrate on this special case here because it simplifies the mathematics, and so facilitates insight into the essential features of the account.

I shall let R denote that method M is completely reliable, so that $\bar{R}$ denotes that M is merely a random method. Then it is part of the intended meaning of R that

$$P(H|RM_H) = 1. \qquad (2)$$

https://doi.org/10.1086/psaprocbienmeetp.1988.1.192994 Published online by Cambridge University Press

It follows from this that $P(H|RM_H E\bar{O}) = P(H|RM_H EO) = 1$; thus if we were sure that M was completely reliable, the predictivist thesis (1) would be false. Except in pathological circumstances, the same would also be true if we were sure that M was a random method, i.e. we would have

$$P(H|\bar{R}M_H E\bar{O}) = P(H|\bar{R}M_H EO). \tag{3}$$

This says that if we are given that the method is random, and that it generated a hypothesis H which was correct about evidence E, then the probability of H is the same whether evidence E was accommodated or predicted. For instance, in the coin-tossing example of the preceding section, if we are given that the subject has only a random chance of getting a prediction right, we would normally say that the probability of the 100th toss agreeing with the subject's prediction is 1/2, regardless of whether the first 99 tosses were accommodated or predicted by the subject, and so (3) is satisfied. On the other hand, one can imagine pathological cases in which (3) would not hold, because O is itself relevant evidence concerning H. For example, the coin tosses might actually be conducted before the experiment is begun; and you might know that the experimenter always runs the experiment involving accommodation of the first 99 tosses when heads occurred on the 100th toss, and otherwise runs the experiment involving prediction of all 100 tosses;[10] then O would be evidence in favor of H, even when we know that the method is random. In such pathological cases, successful prediction can actually be evidence that the asserted hypothesis is false.

This shows that the predictivist thesis does not hold under all possible circumstances; what what we aim to show here is that the thesis holds given certain assumptions which are typically satisfied in science. And (3) will be one assumption of this kind.[11]

From what has been said, it is clear that the predictivist thesis will be true only if we are not completely certain of the reliability of method M; that is, the probability of R must be neither 0 nor 1. And in order for the conditional probabilities which we have just discussed to be definable in the usual way, we need to assume that the conditioning probabilities are non-zero, i.e. we need to assume

$$P(RM_H EO)P(RM_H E\bar{O})P(\bar{R}M_H EO)P(\bar{R}M_H E\bar{O}) > 0. \tag{4}$$

Of course, (4) entails in particular that the probability of R is neither 0 nor 1. I shall assume that (4) holds.

Obviously (1) cannot be true if $P(H|M_H EO) = 1$; thus we will need to assume that $P(H|M_H EO) < 1$. In view of (2) and (4), this assumption is equivalent to

$$P(H|\bar{R}M_H EO) < 1. \tag{5}$$

This assumption would certainly be violated if H were a logical consequence of $\bar{R}M_H EO$; and it might be violated in other cases too. But in the sorts of scientific situations of which our Mendeleyev example is representative, we can expect (5) to be satisfied.

There are a couple of additional assumptions, asserting the irrelevance of O to other propositions of interest, which we shall need, and which we can expect to be satisfied in normal scientific circumstances. One is the following:

$$P(R|E\bar{O}) = P(R|EO). \tag{6}$$

In words: If we have no information about what hypothesis M has generated, but do know that evidence E obtains, then the further information that the truth value of E has

https://doi.org/10.1086/psaprocbienmeetp.1988.1.192994 Published online by Cambridge University Press

been input to M neither increases nor decreases the probability that M is reliable. Thus in the coin-tossing case, if we know the result of the first 99 tosses, but do not know what hypothesis the subject has proposed, then the further information that the subject was given the results of those first 99 tosses before being asked to propose a hypothesis is no reason to think the subject is or is not reliable. This condition could be violated in a pathological situation like the one we considered in discussing (3). It is also clear that if (6) is violated, then the mere fact that prediction has occurred might itself be evidence that the method is unreliable, and thus the predictivist thesis could fail. So we will assume (6).[12]

Let M_E denote that M has generated a hypothesis which entails E. A second assumption about the irrelevance of O, which we shall need, is that

$$\frac{P(M_H|M_E REO)}{P(M_H|M_E \bar{R} EO)} = \frac{P(M_H|M_E RE\bar{O})}{P(M_H|M_E \bar{R} E\bar{O})}. \tag{7}$$

As a simple illustration, consider our coin-tossing example again. Here we could reasonably expect to have

$$P(H|M_E REO) = P(H|M_E RE\bar{O}) = P(H|E). \tag{8}$$

From (2), and the fact that the subject must take a stand on the truth value of H, we also have $P(M_H|HR) = 1$. Applying (2) again, it then follows that $P(M_H|M_E REO)$ and $P(M_H|M_E RE\bar{O})$ are both equal to $P(H|E)$.[13] Also we can suppose that the probability of a random method predicting heads on the 100th toss is $1/2$, in which case $P(M_H|M_E \bar{R} EO) = P(M_H|M_E RE\bar{O}) = 1/2$. It then follows that (7) holds. Note that if the coin is thought to be fair, then $P(H|E) = 1/2$, and the numerator and denominator are the same on both sides. On the other hand, if the coin is thought to be biased for heads, say, then the numerator will be greater than the denominator on each side, but (7) continues to hold.

Scientific cases differ from our coin-tossing example in that scientists, unlike our experimental subject, are permitted to suspend judgment. Thus scientists can use hypothesis-generating methods which may fail to generate any hypothesis on a given topic. Under these circumstances, it is quite consistent with (2) to have $P(M_H|HR) < 1$. Then even if (8) continues to hold, the numerators in (7) need not be equal. Since predicting H is less risky when E is known than when E is not known, we might well have

$$P(M_H|M_E REO) > P(M_H|M_E RE\bar{O}).$$

If this is so, then for (7) to hold we must also have

$$P(M_H|M_E \bar{R} EO) > P(M_H|M_E \bar{R} E\bar{O}),$$

and this seems likely to hold for the same reason.

Of course, it is possible for (7) to be violated; and when it is, the predictivist thesis can also fail. This may be illustrated in our coin-tossing example by supposing we know that if M is random, then almost certainly it predicts unobserved coin tosses by supposing they will conform to a certain specific sequence, and that in this sequence the first element is tails and the 100th element is heads. Then given the first 99 tosses, the method if random will almost certainly predict tails on the 100th toss, but if no tosses are given then the method will almost certainly predict heads on the 100th toss. Then

$$P(M_H|M_E \bar{R} EO) \approx 0; \quad P(M_H|M_E \bar{R} E\bar{O}) \approx 1.$$

https://doi.org/10.1086/psaprocbienmeetp.1988.1.192994 Published online by Cambridge University Press

Assuming (8) continues to hold, we still have $P(M_H|M_EREO) = P(M_H|M_ERE\bar{O}) = P(H|E)$, and therefore have a violation of (7). Also, in this example M_HEO confirms R more strongly than does $M_HE\bar{O}$, and hence it is possible for the predictivist thesis to fail. But situations such as this seem unlikely to arise in normal scientific contexts, and so I shall assume that (7) holds.[14]

Let $M_{\bar{E}}$ denote that M has generated a hypothesis inconsistent with E, and let $M_{E,\bar{E}}$ denote that either M_E or $M_{\bar{E}}$ obtains. A further assumption which we shall need is that

$$P(M_{E,\bar{E}}|RE\bar{O}) \geq P(M_{E,\bar{E}}|\bar{R}E\bar{O}). \tag{9}$$

That is, if the truth-value of E is not input to the method (and if E is true), then the probability of M generating a hypothesis which takes a stance on the truth value of E is at least as great when M is reliable as when M is random. In our coin-tossing example, we have $P(M_{E,\bar{E}}) = 1$, and hence have the special case of (9) in which both sides of the inequality are equal. In Mendeleyev's case, one might not be sure in advance that his method would produce a hypothesis which asserted or denied the existence of an element with the properties of gallium (say). Thus we could have $P(M_{E,\bar{E}}|E\bar{O}) < 1$, and so our reason for thinking that (9) holds in the coin-tossing example does not apply here. But the fact that Mendeleyev was bold enough to make his unprecedented prediction indicates that he was confident of the reliability of his method; and if we take this confidence to be evidence for the actual reliability of his method (or at least not evidence for its unreliability),[15] then we will have

$$P(R|M_{E,\bar{E}}E\bar{O}) > P(R|\bar{M}_{E,\bar{E}}E\bar{O}).$$

It now follows from Bayes' theorem that (9) holds in Mendeleyev's case. On the other hand, if (9) were to fail, so that $P(M_{E,\bar{E}}|\bar{R}E\bar{O}) > P(M_{E,\bar{E}}|RE\bar{O})$, then the mere fact that a prediction of the truth value of E has been made can be evidence for $\bar{R}$, and under these circumstances the predictivist thesis can fail. So I shall assume that (9) holds.

To derive the predictivist thesis, we will also need to assume that if M is random, then there is a positive probability that it would predict $\bar{E}$ when E obtains, i.e.

$$P(M_{\bar{E}}|\bar{R}E\bar{O}) > 0. \tag{10}$$

To see how (10) may fail, let M be the method which predicts that unobserved coin tosses will always land heads, and suppose the coin is fair. Then M has only a random chance of success, although (10) is violated if E is evidence that the coin landed heads on a certain toss. However, in realistic situations we typically do not know enough about a method to be sure what hypothesis it will generate, in which case (10) will hold. This is clearly so when the method is identified merely as "the subject's method of predicting coin tosses". We can similarly interpret "Mendeleyev's method of projecting patterns in the elements" as being sufficiently underspecified that (10) holds.

We are now almost done with assumptions. We just need to note two points. First, we are assuming (again for simplicity) that hypothesis H entails evidence E; consequently, we have

$$P(M_E|M_H) = 1. \tag{11}$$

We also assume that if evidence E is input to method M, then M will generate a hypothesis which entails E; that is,

$$P(M_E|EO) = 1. \tag{12}$$

Of course, evidence is in reality often uncertain, so that it can be reasonable to entertain hypotheses that contradict some available evidence. But we are here making the idealization that the evidence E is certainly true, and that allows us to assume (12) holds.

It remains to show that the assumptions made in this section suffice to entail the predictivist thesis (1).

5. Derivation of the predictivist thesis

According to the intuitive argument sketched in section 3, successful prediction increases the probability that the hypothesis generation method being used is reliable. This intuition can be expressed as

$$P(R|M_H E\bar{O}) > P(R|M_H EO). \tag{13}$$

The argument then goes on to say that (13) in turn entails (1). I will show that both parts of this argument are valid, given the assumptions we have specified.

I begin with the derivation of (13). By the theorem of total probability,

$$P(E|M_{E,\bar{E}}R) = P(E|M_E R)P(M_E|M_{E,\bar{E}}R) + P(E|M_{\bar{E}}R)P(M_{\bar{E}}|M_{E,\bar{E}}R).$$

From (2), with E and $\bar{E}$ substituted for H, we have $P(E|M_E R) = P(\bar{E}|M_{\bar{E}}R) = 1$. Hence

$$P(E|M_{E,\bar{E}}R) = P(M_E|M_{E,\bar{E}}R). \tag{14}$$

Now Bayes' theorem, (14), and (2) with E substituted for H, give us successively:

$$\begin{aligned} P(M_E|M_{E,\bar{E}}RE) &= \frac{P(E|M_E R)P(M_E|M_{E,\bar{E}}R)}{P(E|M_{E,\bar{E}}R)} \\ &= P(E|M_E R) \\ &= 1. \end{aligned} \tag{15}$$

Thus

$$\begin{aligned} 1 &> \frac{P(M_E|M_{E,\bar{E}}\bar{R}E\bar{O})}{P(M_E|M_{E,\bar{E}}RE\bar{O})}, && \text{by (10) and (15)} \\ &= \frac{P(M_E|M_{E,\bar{E}}\bar{R}E\bar{O})P(M_{E,\bar{E}}|\bar{R}E\bar{O})}{P(M_E|M_{E,\bar{E}}RE\bar{O})P(M_{E,\bar{E}}|RE\bar{O})}, && \text{by (9)} \\ &= \frac{P(M_E|\bar{R}E\bar{O})}{P(M_E|RE\bar{O})}, && \text{since } M_E \text{ entails } M_{E,\bar{E}}. \end{aligned}$$

Since (12) implies that $P(M_E|REO) = P(M_E|\bar{R}EO) = 1$, we then have that

$$\frac{P(M_E|\bar{R}EO)}{P(M_E|REO)} > \frac{P(M_E|\bar{R}E\bar{O})}{P(M_E|RE\bar{O})}.$$

So by (7),

$$\frac{P(M_H|M_E\bar{R}EO)P(M_E|\bar{R}EO)}{P(M_H|M_EREO)P(M_E|REO)} > \frac{P(M_H|M_E\bar{R}E\bar{O})P(M_E|\bar{R}E\bar{O})}{P(M_H|M_ERE\bar{O})P(M_E|RE\bar{O})}.$$

In view of (11), it follows that

$$\frac{P(M_H|\bar{R}EO)}{P(M_H|REO)} > \frac{P(M_H|\bar{R}E\bar{O})}{P(M_H|RE\bar{O})}.$$

Together with (4) and (6), this implies that

$$
\begin{aligned}
1 &< \frac{P(R|EO) + P(\bar{R}|EO)\frac{P(M_H|\bar{R}EO)}{P(M_H|REO)}}{P(R|E\bar{O}) + P(\bar{R}|E\bar{O})\frac{P(M_H|\bar{R}E\bar{O})}{P(M_H|RE\bar{O})}}; \\
&= \frac{P(M_H|RE\bar{O})}{P(M_H|REO)} \times \frac{P(R|EO)P(M_H|REO) + P(\bar{R}|EO)P(M_H|\bar{R}EO)}{P(R|E\bar{O})P(M_H|RE\bar{O}) + P(\bar{R}|E\bar{O})P(M_H|\bar{R}E\bar{O})} \\
&= \frac{P(M_H|RE\bar{O})P(M_H|EO)}{P(M_H|REO)P(M_H|E\bar{O})}, \qquad \text{by the theorem of total probability} \\
&= \frac{P(M_H|RE\bar{O})P(R|E\bar{O})P(M_H|EO)}{P(M_H|REO)P(R|EO)P(M_H|E\bar{O})}, \qquad \text{by (4) and (6)} \\
&= \frac{P(R|M_HE\bar{O})}{P(R|M_HEO)}, \qquad \text{by Bayes' theorem.}
\end{aligned}
$$

This completes the proof of (13).

We now proceed to show that (13) entails (1), given our assumptions. By the theorem of total probability,

$$\frac{P(H|M_HE\bar{O})}{P(H|M_HEO)} = \frac{P(R|M_HE\bar{O})P(H|RM_HE\bar{O}) + P(\bar{R}|M_HE\bar{O})P(H|\bar{R}M_HE\bar{O})}{P(R|M_HEO)P(H|RM_HEO) + P(\bar{R}|M_HEO)P(H|\bar{R}M_HEO)}.$$

From (2) we have that $P(H|RM_HE\bar{O}) = P(H|RM_HEO) = 1$; and by (3) we have that $P(H|\bar{R}M_HE\bar{O}) = P(H|\bar{R}M_HEO) = \delta$, say. Thus

$$
\begin{aligned}
\frac{P(H|M_HE\bar{O})}{P(H|M_HEO)} &= \frac{P(R|M_HE\bar{O}) + \delta P(\bar{R}|M_HE\bar{O})}{P(R|M_HEO) + \delta P(\bar{R}|M_HEO)} \\
&= \frac{(1-\delta)P(R|M_HE\bar{O}) + \delta}{(1-\delta)P(R|M_HEO) + \delta}.
\end{aligned}
$$

Assumption (5) asserts that $\delta < 1$; consequently, the above equation together with (13) entails (1).

6. Concluding remarks

We have now shown that Bayesians can agree with the predictivist thesis, and further can offer a Bayesian explanation of why it holds when it does. In the course of developing this explanation, we made a sizable number of assumptions. A few of these were idealizations designed to simplify the formal analysis; for example, the assumption that the hypothesis generation method is either completely reliable or else random. I plan to show on another occasion that it is possible to relax these assumptions, and still give essentially the same explanation of the predictivist thesis. However, most of the assumptions we made were found to be necessary, in that without them the predictivist thesis could fail. This category includes such assumptions as that we not be certain of the reliability of the method, and that the occurrence of accommodation should not itself confirm the reliability of the method. It is because these latter assumptions do normally hold in scientific contexts that the predictivist thesis is true in those contexts.

Those who have accepted the predictivist thesis seem not to have been aware that the thesis can fail under certain conditions. I see it as one of the merits of the present account that it reveals this fact.

https://doi.org/10.1086/psaprocbienmeetp.1988.1.192994 Published online by Cambridge University Press

At least until fairly recently, it has been generally accepted in the philosophy of science that the method by which a hypothesis was discovered is irrelevant to the confirmation (or corroboration) of that hypothesis (Popper 1959, p. 31f.). A related view is that there can be no "logic of discovery", i.e. no way of identifying discovery methods which are more likely than others to yield true hypotheses. A corollary of our explanation of the predictivist thesis is that these widely accepted views are incorrect. For we have seen that successful prediction provides reason to think that a discovery method is reliable; and since reason to believe a method reliable is reason to believe the hypotheses it generates are true, it follows that the method by which a hypothesis is generated is indeed relevant to the confirmation of that hypothesis.[16]

According to the account I have offered, the confirmatory relevance of discovery methods is central to understanding the predictivist thesis. The fact that this relevance has been widely denied probably explains why the explanation of the predictivist thesis has eluded so many eminent philosophers of science.

Notes

[1]This paper was written during my tenure as a fellow with the Michigan Society of Fellows. It is based on research supported by the National Science Foundation under grant SES-8708168.

[2]I am using concatenation to represent conjunction, and overbars to represent negation. Thus $E\bar{O}$ is the proposition that E is true and O is false.

[3]Horwich does offer an analogy which is intended to support the statement I have just quoted. In this analogy, the statements considered are:

F = my car is green
H_P = the previous owner painted the car an ineradicable green
H_I = I insisted upon a green car.

Horwich claims that "the only element in $[H_I]$ which is relevant to the probability of H_P is that my car is green", and thus concludes that

$$P(H_P|H_IF) = P(H_P|F).$$

The anti-predictivist position is then supposed to be analogous to this. But Horwich's account of his example is faulty. To see this, let

F_B = my car was green when I bought it.

By the theorem of total probability,

$$P(H_P|H_IF) = P(H_P|H_IFF_B)P(F_B|H_IF) + P(H_P|H_IF\bar{F}_B)P(\bar{F}_B|H_IF).$$

We can plausibly suppose that the relevance of H_I to H_P is fully accounted for by the fact that H_I entails F_B, in which case $P(H_P|H_IFF_B) = P(H_P|FF_B)$. Also, since H_I entails F_B, $P(F_B|H_IF) = 1$ and $P(\bar{F}_B|H_IF) = 0$. Thus we obtain

$$P(H_P|H_IF) = P(H_P|FF_B).$$

Now the theorem of total probability also gives us

$$P(H_P|F) = P(H_P|FF_B)P(F_B|F) + P(H_P|F\bar{F}_B)P(\bar{F}_B|F).$$

Since $P(H_P|F\bar{F}_B) = 0$, we then have

$$P(H_P|F) = P(H_P|FF_B)P(F_B|F).$$

Now the fact that my car is green does not imply that it was green when I bought it, and so it is reasonable to suppose that $P(F_B|F) < 1$, in which case it follows that $P(H_P|H_IF) > P(H_P|F)$, contrary to Horwich's claim. So if there is any analogy here at all, it is one that supports the predictivist thesis.

[4]For the history of the development of the periodic table, see Ihde (1964), chapter 9.

[5]Though there was doubt about the correctness of the reported properties for some of these elements, and Mendeleyev put question marks against five of the entries in the table (which is reproduced in Ihde (1964, p. 245).

[6]Campbell and Vinci (1982, 1983) have made two attempts to give a Bayesian explanation of a variant form of the predictivist thesis. What they argue for is that successful prediction confirms the approximate truth of a hypothesis more than does successful accommodation. (Though they prefer to speak of "heuristically novel" evidence, rather than predicted evidence.) Unfortunately, neither of these attempts is successful. In the 1982 paper, they derive their conclusion by assuming, amongst other things, that (i) the probability of a hypothesis H being close to the truth, given only that it has successfully accommodated evidence E, must be very small; and (ii) the probability of H fitting evidence E, given that H is not close to the truth and was not designed to accommodate E, is vanishingly small. But these assumptions beg the question against the anti-predictivist, who will hold that when H and E are such that (ii) is satisfied, then (i) is false. As for the 1983 paper, let N be the proposition that of the hypotheses generated by method M which are not close to the truth, very few entail E; also let T be the proposition that H is close to the truth. Then what Campbell and Vinci establish in this second paper is that (in my notation):

$$P(T|M_HEN\bar{O}) > P(T|M_HEO).$$

But here the extra confirmation on the left-hand side comes from N, not from $\bar{O}$; and so this result fails to establish that prediction has any special confirmatory value.

[7]J.J. Berzelius, *Lehrbuch der Chemie* (1845); cited in Ihde (1964), p. 237.

[8]In particle physics, the method led to the discovery of the "eightfold way" by Gell-Mann and Ne'eman (1964).

[9]In the latter case, the probability will depend on the proportion of true statements in the class from which the hypothesis was chosen.

[10]"Prediction" is now the description by the subject of tosses whose outcomes have not been revealed to the subject, though they are known to the experimenter.

[11]Actually, we do not need the full strength of (3); for the arguments which follow, it would suffice to assume the weaker condition $P(H|\bar{R}M_HE\bar{O}) \geq P(H|\bar{R}M_HEO)$.

[12]As this discussion suggests, it would suffice for our purposes to make the weaker assumption that $P(R|E\bar{O}) \geq P(R|EO)$.

[13]Proof of the last two statements: Since the subject must take a stance on the truth value of H, $\bar{M}_H$ implies that M has generated a hypothesis inconsistent with H, whence it follows from (2) that $P(H|\bar{M}_H R) = 0$. It then follows from Bayes' theorem that $P(\bar{M}_H|HR) = 0$, and hence that $P(M_H|HR) = 1$. Also (2) and Bayes' theorem entail that $P(M_H|\bar{H}R) = 0$. Now by the theorem of total probability,

$$\begin{aligned} P(M_H|M_E REO) &= P(M_H|HM_E REO)P(H|M_E REO) + \\ &\qquad P(M_H|\bar{H}M_E REO)P(\bar{H}|M_E REO) \\ &= P(H|M_E REO), \qquad \text{by the above results} \\ &= P(H|E), \qquad \text{by (8).} \end{aligned}$$

A similar proof shows that $P(M_H|M_E RE\bar{O})$ is also equal to $P(H|E)$.

[14]Again, a weaker assumption would suffice; in this case, it would be enough to assume that

$$\frac{P(M_H|M_E REO)}{P(M_H|M_E \bar{R}EO)} \leq \frac{P(M_H|M_E RE\bar{O})}{P(M_H|M_E \bar{R}E\bar{O})}.$$

[15]Let C be the event that Mendeleyev is confident of the reliability of his method. Then the first assumption stated here can be represented more precisely as

$$P(C|M_{E,E}E\bar{O}) > P(C|\bar{M}_{E,E}E\bar{O}).$$

The second assumption is that

$$P(R|CM_{E,E}E\bar{O}) \geq P(R|\bar{C}M_{E,E}E\bar{O}), \quad \text{and} \quad P(R|C\bar{M}_{E,E}E\bar{O}) \geq P(R|\bar{C}\bar{M}_{E,E}E\bar{O}).$$

References

Campbell, R. and Vinci, T. (1982), "Why are Novel Predictions Important?", *Pacific Philosophical Quarterly 63*: 111–121.

------------------- (1983), "Novel Confirmation", *British Journal for the Philosophy of Science 34*: 315–341.

Duhem, P. (1914), *La Théorie Physique: Son Objet, Sa Structure*. Second edition. Paris: Marcel Rivière & Cie. Translated by P.P. Wiener as *The Aim and Structure of Physical Theory*, Princeton: Princeton University Press (1954).

Gardner, M.R. (1982), "Predicting Novel Facts", *British Journal for the Philosophy of Science 33*: 1–15.

Gell-Mann, M. and Ne'eman, Y. (eds.) (1964), *The Eightfold Way*. New York and Amsterdam: W.A. Benjamin.

Hempel, C.G. (1966), *Philosophy of Natural Science*. Englewood Cliffs: Prentice-Hall.

Horwich, P. (1982), *Probability and Evidence*. Cambridge: Cambridge University Press.

Huygens, C. (1690), *Traite de la Lumiere*. Leiden: Pierre vander Aa. Translated by S.P. Thompson as *Treatise on Light*, London: Macmillan (1912).

Ihde, A.J. (1964), *The Development of Modern Chemistry*. New York: Harper and Row.

Kelly, K.T. (1987), "The Logic of Discovery", *Philosophy of Science 54*: 435–452.

Keynes, J.M. (1921), *A Treatise on Probability*. London: Macmillan.

Kuhn, T.S. (1977), *The Essential Tension*. Chicago: University of Chicago Press.

Lakatos, I. (1970), "The Methodology of Scientific Research Programmes", in (Lakatos and Musgrave 1970, pp. 91–196).

Lakatos, I. and Musgrave, A. (eds.) (1970), *Criticism and the Growth of Knowledge*. London: Cambridge University Press.

Leibniz, G.W. (1678), "Letter to Herman Conring", in (Leibniz 1969, pp. 186–191).

---------- (1969), *Philosophical Papers and Letters*, ed. L.E. Loemker. Dordrecht: D. Reidel.

Mill, J.S. (1872), *A System of Logic*. Eighth edition. London: Longmans, Green and Co..

Peirce, C.S. (1883), "A Theory of Probable Inference", in (Peirce 1883, pp. 126–181).

-------- (ed.) (1883), *Studies in Logic*. Boston: Little, Brown.

Popper, K.R. (1959), *The Logic of Scientific Discovery*. New York: Harper and Row.

--------- (1965), *Conjectures and Refutations*. Second edition. New York: Harper and Row.

Rosenkrantz, R.D. (1977), *Inference, Method and Decision*. Dordrecht: D. Reidel.

Schlesinger, G.N. (1987), "Accommodation and Prediction", *Australasian Journal of Philosophy 65*: 33–42.

Simon, H.A. (1955), "Prediction and Hindsight as Confirmatory Evidence", *Philosophy of Science 22*: 227–230. Reprinted in Simon (1977).

--------- (1977), *Models of Discovery*. Dordrecht: D. Reidel.

Whewell, W. (1847), *The Philosophy of the Inductive Sciences*. Second edition. London: Parker.