Brit. J. Phil. Sci. **42** (1991), 547–555 *Printed in Great Britain*

DISCUSSION

The 'Old Evidence' Problem

ABSTRACT

This paper offers an answer to Glymour's 'old evidence' problem for Bayesian confirmation theory, and assesses some of the objections, in particular those recently aired by Chihara, that have been brought against that answer. The paper argues that these objections are easily dissolved, and goes on to show how the answer it proposes yields an intuitively satisfactory analysis of a problem recently discussed by Maher. Garber's, Niiniluoto's and others' quite different answer to Glymour's problem is considered and rejected, and the paper concludes with some brief reflections on the prediction/accommodation issue.

1 INTRODUCTION: GLYMOUR'S POSER

Clark Glymour [1980] presented subjective Bayesians with the following problem. Suppose a hypothesis *h* is proposed which turns out to explain some already well-known data *e*. Can *h* ever be confirmed by *e*? Intuitively, the answer is that it can. Examples abound, in fact, of *h*s apparently well-supported by such *e*s. One very often mentioned is General Relativity with its allegedly strong support from the data on the annual precession of Mercury's perihelion, data already fifty years old when the field equations of General Relativity were first obtained by Einstein.

The problem for subjective Bayesians is explaining and justifying the attribution of such support. For Bayes's Theorem says that

$$P(h/e) = \frac{P(e/h)P(h)}{P(e)} \tag{1}$$

and the received wisdom is that for subjective Bayesians these probabilities are all relativized to the individual's stock K of contemporary background information. This relativization has the following unpleasant consequence, however. If e is known at the time h is proposed, then e is in K and so $P(e)=P(e/h)=1$, giving, from (1), $P(h/e)=P(h)$; which means that e gives no support to h.

2 AND AN ANSWER

One answer—and I think the correct one—to Glymour's nasty problem, often called the *Old Evidence problem*, is to deny that when assessing support according to the difference between $P(h/e)$ and $P(h)$, the probabilities should be relativized to K; rather they should *always* be relativized to $K-\{e\}$ (I shall discuss the question of whether this quantity can be consistently defined in a moment). And why? The answer is straightforward. When you ask yourself how much support e gives h, you are plausibly asking how much a knowledge of e *would* increase the credibility of h, which is the same thing as asking how much e boosts the credibility of h relative to what *else* you currently know. The 'what else' is just $K-\{e\}$.

Glymour considered this way of evaluating support in the case in which e is already known, but not favourably. His objection is a curious one, however: it seems to be merely that there is no general procedure for computing $P(e)$ in such cases (p. 87). But the subjective Bayesian will not be disturbed by this observation, because $P(e)$ and $P(h)$ are explicitly regarded as exogenous parameters within his theory, whatever they're relativized to. It may take some exercise of the imagination to evaluate what your degree of belief in h would be were you, counterfactually, not to know e, but there is no reason in principle to think that it can't be done. Indeed, it seems to be done all the time: 'if I hadn't seen him palm that card, I'd think he had paranormal powers', for example.

3 CHIHARA'S OBJECTIONS

While Glymour's objection seems easily answered, it is not the only one to have been brought against the $K-\{e\}$ relativization for the probabilities used in computing support. Chihara's recent objections [1987] turn on the exact nature of $K-\{e\}$. His first objection is as follows: if $K-\{e\}$ signifies merely the set-theoretic subtraction of e from K, and K itself is intended always to be a deductively closed set of sentences, then $K-\{e\}$ will not be well defined if there are other sentences in K entailing e. Nor is there any unambiguous direction concerning which to remove if there are. For if a and b jointly but not separately entail e then there is a choice of whether to remove a or b or both; but that choice will be arbitrary. This is not the only problem. If K is deductively closed and e is simply deleted from K then all the consequences of e

by itself and jointly with the remainder of K will remain. This means that e will almost certainly be very probable relative to $K-\{e\}$, which may not be as serious as assigning e invariably probability one, but is not much less so.

I think that it is possible to answer these objections, and the answer is twofold. It is, firstly, to drop the insistence on the deductive closure of K, or indeed of any body of background information, and secondly, to regard K as, in effect, an independent axiomatization of background information and $K-\{e\}$ as the simple set-theoretic subtraction of e from K. The requirement of deductive closure is quite unnecessary anyway within a Bayesian theory, since the probability function P_A relativized to any body A of information will necessarily assign the value 1 to any deductive consequence of the sentences in A whether or not that consequence itself is explicitly included in A (note that P_A assigns 1 because the notion of tautology is widened in effect to include consequences of A, not because $P_A(\cdot)$ is regarded—it is not—as a conditional probability $P(\cdot/A)$). Moreover, one's representation of one's background information is never a deductively closed set, and is typically a set of logically discrete statements, insofar as it is articulated at all. So modelling background information in this way, while an idealization, is not outrageously unverisimilar.

Suppose, though, that e is not an isolated part of background information: in other words it does not appear explicitly as a member of K where K is now characterized in the way I suggest. Suppose, for example, with Chihara, that e is nevertheless a strict consequence of some of the sentences in K. It may even be the case that among these other sentences in K were theories initially contrived for the explanation of e, which eventually became so inductively grounded that they acquired the status of certainty. What then? Here there seems to be no determinate way of characterizing one's state of knowledge on the counterfactual assumption that one does not know e, since not knowing e plausibly entails—if one knows the logical implications between the sentences in K and e—not knowing at least one other item of information currently in K (I am investing the concept of knowledge here with only a very minimal content, incidentally: namely, that if you know that $\{a_1, \ldots, a_n\}$ implies b, and you know $a_1, \ldots, a_n$, then you know b). But that is all that can be said, and we must conclude that in such cases the probabilities in the support function are relativized to an indeterminate state of background information, with the consequence that the support e gives to any newly advanced hypothesis is also indeterminate. However, the sorts of cases which are brought up in the literature tend to be those in which the evidence, like the statement describing the magnitude of the observed annual advance of Mercury's perihelion, is a logically isolated component of background information; and here relativizing to the set-theoretical subtraction of e from an—in effect—axiomatized K is well-defined.

But there appears to be a very powerful and familiar objection to regarding K as an axiomatization of background knowledge. This is that a deductive theory can be axiomatized in different ways, and the effect of removing one and the same axiom from two equivalent axiomatizations may well give rise to different consequence classes *i.e.* to non-equivalent theories. A simple example of this is the pair $\{a, b\}$ and $\{a, a \to b\}$ where a and b are logically independent sentences. Clearly, a and $a \to b$ are also independent and the axiom pairs generate the same sets of consequences, but if we remove a from both the resulting sets of consequences are distinct: b is in one and not the other, for example.

The representation-dependence of 'background minus e' should not, however, disturb us, disturbing though it may well appear at first sight. It looks disturbing only because we have inherited an ancient habit of thought that dies hard, namely that there is, somewhere in logical space, a purely objective logic of induction. There is not. There are sound planks that float amid the wreck of this beguiling idea, however; they constitute a theory of coherent degrees of belief, measured by a probability function, and a theory of rather weakly constrained inference based upon the properties of a *conditional* probability function. The resulting theory of belief-cum-inference is called Personalistic or Subjective Bayesianism. Now it is entirely consistent with this theory that different individuals may structure their knowledge in different ways. And if they do, it is only to be expected that the effect of selectively deleting items from this stock will vary from individual to individual. In that case, people's estimates of the confirmatory effect of old evidence may in principle vary (though shared educational experience means that it probably would not).

Let us now address Chihara's second objection to relativizing probabilities to $K-\{e\}$ is that while e may be logically isolated in this way, it may not be probabilistically isolated from the remainder of K ([1987] p. 553). Chihara points out that there may be things in K that are there because of the same observations that engendered e, and that these may make e itself highly probable: *e.g.* that the highly competent researcher A believes e may be a consequence of those observations; and this last statement makes e itself probable. But it is difficult to see why Chihara regards this as an objection. If e is still very probable relative to $K-\{e\}$, then all this means is that it will not count as a strong confirmation of h. There is nothing problematic about this: on the contrary.

4 HOW THE ANSWER SOLVES A PROBLEM

The extant objections to relativizing probabilites to $K-\{e\}$ in computing supports do not appear to be very difficult to answer. As we have seen, such a strategy is natural enough when we analyse incremental support in terms of the increase in credibility of h which *would now* be contributed by a knowledge of e. Moreover, the relativization to $K-\{e\}$ does seem to be essentially involved if we are to make sense of our intuitive evaluations of support. A rather striking example, over and above examples of the Mercury perihelion sort, is afforded by the following situation envisaged recently by Maher ([1988]), and further discussed in Howson ([1989b]) and Howson and Franklin ([1990]). We are asked to consider two possible states of affairs:

(i) an individual predicts the outcomes of the next 100 tosses of a coin. The coin is tossed 99 times, and the outcomes of these tosses are as predicted. This fact gives us a very high degree of confidence in the prediction of the outcome of the 100th toss;

(ii) an individual predicts the same outcome of the 100th toss as in (i), but now *after* having learned the outcomes of the first 99 tosses. Let us suppose that roughly sixty per cent of the outcomes of the 99 tosses in both this case and in (i) are heads, and that the predicted outcome of the 100th toss is of a head. Our intuitive degree of confidence in this prediction of the 100th toss, measured by the probability we assign that outcome in the light of all the data, is here probably around sixty per cent, whereas in (i) it is practically one.

Now for a Bayesian analysis of these judgments of support. First, some abbreviations. Let h be the hypothesis that a head will occur at the 100th toss of the coin, and let e be the description of the outcomes of the 99 tosses. A hypothesis whose consideration is clearly crucial for understanding our intuitive verdicts is that the subject has prior access, by some unspecified means, to reliable information about the outcomes of the 100 tosses. In (i) this hypothesis, call it m, is well supported by e, and in (ii) it is not. That is really the nub of the issue, and we shall now see that the relativization to $K-\{e\}$ is essential to its explanation.

By the probability calculus

$$P(h/e)=P(h/m\&e)P(m/e)+P(h/-m\&e)P(-m/e) \qquad (2)$$

In both (i) and (ii) $P(h/m\&e)$ is equal to 1, since we know that the subject predicted h, and therefore m implies that h is true. Hence in (i) and (ii)

$$P(h/e)=P(m/e)+P(h/-m\&e)P(-m/e) \qquad (3)$$

Now $P(m/e)=P(e/m)P(m)$ divided by $P(e)$. Let us now consider case (i). Here $P(e/m)=1$, again since it is background information that the subject predicted e as well as h (though it is not in background information that e is true), and so m implies e modulo this information. Moreover,

$$P(e)=P(e/m)P(m)+P(e/-m)P(-m),$$

and $P(e/m)=1$ while $P(e/-m)$ is going to be very small indeed: the latter quantity is really just the probability of predicting correctly 99 tosses of a coin by chance. Hence $P(e)$ is approximately equal to $P(m)$ and so $P(m/e)$ is close to 1; therefore so too is $P(h/e)$.

Now for (ii). What is $P(m/e)$ here? Well, the background information in this case now includes the information that e is true, and also that the subject predicted h after being acquainted with e's truth. Let us look, as in (i), at $P(e/m)$ and $P(e)$, where, in the light of the discussion in the Section 3, we are relativizing the probabilities to background minus e. Relative to this truncated background, there is no more reason to believe in the truth of e given m than otherwise, and so we can plausibly (*i.e.* following the reasoning processes of *l'homme moyen sensuel*, whom in this context we can take to possess an averagely unbiased outlook on this sort of problem at least) set $P(e/m)$ equal to $P(e)$. Hence $P(m/e)=P(m)$, which, we can assume, is extremely small. Hence we can equate $P(h/e)$ approximately to $P(h/-m\&e)$. Taking $-m$ as before to be equivalent to the null hypothesis of a purely chance agreement between prediction and outcome, we obtain, by the usual Bayesian reasoning, a value for $P(h/e)$ reflecting (approximately) the proportion of heads in e.

So we explain our intuitive verdicts by relativizing to $K-\{e\}$ in both cases, though in (i) e was not in K and so the relativization made no difference. At this point it might be objected that we could more easily have explained the difference in support for h in (i) and (ii) by setting $P(e/h)=P(e)=1$ in (ii), in line with Glymour's prescription. But the objection cannot be sustained, because it is easy to see that were we to follow that strategy we should ignore the constitution of the sample e itself in (ii), and yet, as we saw, the measured frequency of heads in that sample is, after all, the sole determinant of our confidence in h. No; the only correct way of explaining our intuitive verdict in (ii), as in the case of Mercury's perihelion, is to relativize to $K-\{e\}$. Lest there be any lingering doubt that Glymour's thesis, that $P(e)$ in the support computation must be set equal to 1 when you learn e's truth, is wrong, consider the classical Bayesian analysis of the hypothesis h' that a coin has a given propensity to yield heads, proposed after noting the coin's mass-distribution. The coin is then tossed n times, and the outcomes described by e'. Does the fact that the experiment delivering e' has been performed mean that $P(e')$ is to be set equal to 1 when $P(h'/e')$ is computed? Of course not; any Bayesian will point out that this is a misuse of the formula.

5 GARBER'S ANSWER

It remains finally to review a quite different attempt to solve the old evidence problem, due to Garber ([1983]), Niiniluoto ([1983]), Jeffrey ([1983]) and others. I shall limit myself to a brief description of Garber's account, though the idea common to all three (and which was first suggested by Glymour) is that it is not the knowledge of the old e which increases the probability of h above its initial level, but learning *that* h *explains* e. For example, Einstein discovered only after writing down the equations of General Relativity that they entailed the anomalous perihelion advance of Mercury. In order to accommodate this sort of logical learning Garber proposes, firstly, extending some initial domain of P to include all statements of the form $h \vdash e$ for all h, e in the initial domain and closing off again under truth functional composition (the entailment sign is treated as a primitive), and secondly, amending the probability axioms by assigning probability 1 only to those *truth-functionally valid* combinations of sentences in the enlarged domain, and adding as a further axiom a 'modus ponens' condition

$$P(a\&b\&(a \vdash b)) = P(a\&(a \vdash b)).$$

The result system models a type of logical as well as empirical learning, at any rate to the extent that it can be proved that there is an infinite number of probability functions P with the property that there are sentences h and e such that $P(e)=1$, $0<P(h \vdash e)<1$, $0<P(h)<1$, and $P(h/h \vdash e>P(h)$.

It does not, of course, follow from this result that any particular case, like the Mercury perihelion one, will exemplify such 'logical learning'; indeed, it is currently far from clear that Garber's analysis can be made to fit all or any such cases. Also, advocates of this whole approach accept uncritically what is has been the burden of the earlier part of this paper to deny, that e's being known must be reflected in its being assigned probability 1. But whatever the virtues and vices of Garber's construction (and it has stimulated a good deal of discussion), *it does not solve the old evidence problem*: it is not, I am now going to argue, the learning merely of an entailment which is responsible for the increase in confirmation when we judge that old evidence supports a new theory. In the first place, we should note that the Garber system is not as it stands valid for more general forms of explanation, like statistical explanation, than entailment modulo satisfaction of initial conditions, yet there seems no reason to suppose that learning of these more general explanatory relations should not exercise the same confirmatory effect on this account. But the logic of explanation in general, unlike that of deduction, is notoriously ill-understood, and incorporating criteria of explanation into the Bayesian system, at any level, is something to be undertaken only, I should have thought, as a very last resort. Secondly, and more important, it is not difficult to think of examples of old evidence e and hypotheses h such that it was already *known* that h explains e, and yet in which e is nevertheless considered to have supported h. Newton was perfectly well aware that Kepler's Laws were explained by his theory when he constructed it, since that this should be so was of course an explicit constraint on the construction of that theory; yet Newton and practically everybody else regarded the inverse square law together with the laws of motion to be supported by Kepler's Laws (Franklin and Howson [1985]). And Earman, in his systematic discussion ([1989]) of the old evidence problem, has pointed out that while Einstein himself may have discovered that his equations entailed the observed value of Mercury's perihelion advance, we still

> 'want to say that the perihelion phenomenon was and is good evidence for Einstein's theory. But along with most students of general relativity, the first thing we may have learned about the theory, even before hearing any details of the theory itself, was that it explains the perihelion advance.' (p. 16.)

Much more homely counterexamples to Garber *et al.*'s thesis can easily be constructed. For example, suppose a coin is tossed n times and the observed relative frequency of heads is r/n. The hypothesis that the coin has a propensity $(r/n) \pm f(n)$ to fall heads is then proposed. According to both statistical practice and theory such a hypothesis is well-supported by the data (in combination with suitable background information), even though it was put forward expressly as the best explanation of that data. The sample-average statistic r/n is what is called a maximum-likelihood estimator of the propensity, and maximum-likelihood estimation is no less than a general algorithm for generating best explanatory hypotheses from the data.

6 CONCLUSION: ACCOMMODATION AND PREDICTION

We have, not entirely unwittingly, strayed into another highly controversial debate, about whether hypotheses deliberately constructed to explain data are thereby at all supported by those data. I have recently argued at length (Howson [1989a], [1989b], and Howson and Urbach [1989] Chapter 11) that hypotheses may in suitable circumstances—which I set out—be confirmed by the data they were constructed to explain, and I have also shown, I believe, where the usual arguments that this is not possible break down. So I hope I shall be permitted to refer to the reader to those sources and not argue any further here for this position. But if, as I hope, it is granted, then it indicates that Garber *et al.*'s diagnosis of why old evidence confirms is incorrect, and—fortunately—we don't have to tamper with classical Bayesianism in the way they recommend.

COLIN HOWSON
London School of Economics

REFERENCES

CHIHARA, C. S. [1987]: 'Some Problems for Bayesian Confirmation Theory', *The British Journal for the Philosophy of Science*, **38**, pp. 551–60.

EARMAN, J. [1989]: 'Old Evidence, New Theories: Two Unresolved Problems in Bayesian Confirmation Theory' (manuscript).

GARBER, D. [1983]: 'Old Evidence and Logical Omniscience', *Testing Scientific Theories*, in J. Earman (*ed*), *Minnesota Studies in the Philosophy of Science*, Vol X, Minneapolis: University of Minnesota Press.

FRANKLIN, A. and HOWSON, C. [1985]: 'Newton and Kepler; a Bayesian Approach', *Studies in History and Philosophy of Science*, **16**, pp. 379–85.

GLYMOUR, C. [1980]: *Theory and Evidence*, Princeton: Princeton University Press.

HOWSON, C. [1985]: 'Some Recent Objections to the Bayesian Theory of Support', *British Journal for the Philosophy of Science*, **36**, pp. 305–9.

HOWSON, C. [1989a]: 'Fitting Your Theory to the Facts: Probably not such a Bad Thing After All', *Discovery, Justification and Evolution of Scientific Theories*, in C. Wade Savage (*ed*), *Minnesota Studies in the Philosophy of Science*, Vol. XIV, Minneapolis: University of Minnesota Press.

HOWSON, C. [1989b]: 'Accommodation, Prediction and Bayesian Confirmation Theory', *PSA 1988*, Vol. 2, Pittsburgh: Pittsburgh University Press.

HOWSON, C. and FRANKLIN, A. [1990]: 'Maher, Mendeleev and Bayesianism', *Philosophy of Science* (forthcoming).

HOWSON, C. and URBACH, P. M. [1989]: *Scientific Reasoning: the Bayesian Approach*, La Salle: Open Court Publishing Company.

JEFFREY, R. C. [1983]: 'Bayesianism with a Human Face', *Testing Scientific Theories*, in J. Earman (*ed*), *Minnesota Studies in the Philosophy of Science*, Vol. X, Minneapolis: University of Minnesota Press.

MAHER, P. [1988]: 'Prediction, Accommodation and the Logic of Discovery', *PSA 1988*, Vol. 1, Pittsburgh: Pittsburgh University Press.

NIINILUOTO, I. [1983]: Novel Facts and Bayesianism', *The British Journal for the Philosophy of Science*, **34**, pp. 375–9.