If we now let $j_n(\omega)$ stand for the number of heads in the first n trials of ω, the weak and strong form of the law of large numbers can be stated as follows:

WLLN The $\mathscr{P}\!r$ measure of the set of ω's for which $|(j_n(\omega)/n) - p| > \varepsilon$ approaches 0 as $n \to \infty$ for any $\varepsilon > 0$.

SLLN The $\mathscr{P}\!r$ measure of the set of all ω's such that $\lim_{n\to\infty}(j_n(\omega)/n) \neq p$ is 0.

To put (SLLN) in its positive form, the Pr probability is one that the limiting relative frequency of heads converges to p.

As indicated in section 7, a form of the weak law of large numbers can be formulated and proved without the help of countable additivity. Roughly, for any $\varepsilon > 0$, the probability (in the objective sense or in the degree-of-belief sense tempered by Lewis's principal principle) that the actually observed relative frequency of heads differs from p by more than ε goes to 0 as the number of flips goes to infinity. This form of the law of large numbers is to be found in the work of Bernoulli. The strong form of the law of large numbers, which requires countable additivity, was not proved until this century (see Billingsley 1979 for a proof).

3 Success Stories

The successes of the Bayesian approach to confirmation fall into two categories. First, there are the successes of Bayesianism in illuminating the virtues and pitfalls of various approaches to confirmation theory by providing a Bayesian rationale for what are regarded as sound methodological procedures and by revealing the infirmities of what are acknowledged as unsound procedures. The present chapter reviews some of these explanatory successes. Second, there are the successes in meeting a number of objections that have been hurled against Bayesianism. The following chapter discusses several of these successful defenses. Taken together, the combined success stories help to explain why many Bayesians display the confident complacency of true believers. Chapters 5 to 9 will challenge this complacency. But before turning to the challenges, let us give Bayesianism its due.

1 Qualitative Confirmation: The Hypotheticodeductive Method

When Carl Hempel published his seminal "Studies in the Logic of Confirmation" (1945), he saw his essay as a contribution to the logical empiricists' program of creating an inductive logic that would parallel and complement deductive logic. The program, he thought, was best carried out in three stages: the first stage would provide an explication of the qualitative concept of confirmation (as in 'E confirms H'); the second stage would tackle the comparative concept (as in 'E confirms H more than E' confirms H''); and the final stage would concern the quantitative concept (as in 'E confirms H to degree r'). In hindsight it seems clear (at least to Bayesians) that it is best to proceed the other way around: start with the quantitative concept and use it to analyze the comparative and qualitative notions. The difficulties inherent in Hempel's own account of qualitative confirmation will be studied in section 2. This section will be devoted to the more venerable hypotheticodeductive (HD) method.

The basic idea of HD methodology is deceptively simple. From the hypothesis H at issue and accepted background knowledge K, one deduces a consequence E that can be checked by observation or experiment. If Nature affirms that E is indeed the case, then H is said to be HD-confirmed, while if Nature affirms $\neg E$, H is said to be HD-disconfirmed. The critics of HD have so battered this account of theory testing that it would be unseemly to administer any further whipping to what is very nearly a dead horse.[1] Rather, I will review the results of the jolly Bayesian postmortem.

Suppose that (a) $\{H, K\} \models E$, (b) $0 < \Pr(H/K) < 1$, and (c) $0 < \Pr(E/K) < 1$.[2] Condition (a) is just the basic HD requirement for confirmation. Condition (b) says that on the basis of background knowledge K, H is not known to be almost surely true or to be almost surely false, and (c) says likewise for E. By Bayes's theorem and (a), it follows that

$$\Pr(H/E \,\&\, K) = \Pr(H/K)/\Pr(E/K). \tag{3.1}$$

By applying (b) and (c) to (3.1), we can conclude that $\Pr(H/E \,\&\, K) > \Pr(H/K)$, i.e., *E incrementally confirms H* relative to K. Thus Bayesianism is able to winnow a valid kernel of the HD method from its chaff.

(To digress, this alleged success story might be questioned on the grounds that HD testing typically satisfies not condition (a) but rather a condition Hempel calls the "prediction criterion" of confirmation; namely, (a′) E is logically equivalent to $E_1 \,\&\, E_2$, $\{H, K, E_1\} \models E_2$, but $\{H, K\} \not\models E_2$. That is, HD condition (a) is satisfied with respect to the conditional prediction $E_1 \rightarrow E_2$, but the total evidence consists of E_1 and E_2 together. Let us use Bayes's theorem to draw out the consequences of (a′). It follows that $\Pr(H/E_1 \,\&\, E_2 \,\&\, K) = \Pr(H/E_1 \,\&\, K)/\Pr(E_2/E_1 \,\&\, K)$. Thus if $\Pr(E_2/E_1 \,\&\, K) < 1$ and $\Pr(H/E_1 \,\&\, K) = \Pr(H/K)$, the total evidence $E_1 \,\&\, E_2$ incrementally confirms H. These latter two conditions are satisfied in typical cases of HD testing. For example, let H be Newton's theory of planetary motion, let E_1 be the statement that a telescope is pointed in such and such a direction tomorrow at 3:00 P.M., and let E_2 be the statement that Mars will be seen through the telescope. Presumably, E_1 is probabilistically irrelevant to the theory, and E_2 is uncertain on the basis of E_1 and K.)

Notice also that from (3.1) it follows that the smaller the value of the prior likelihood $\Pr(E/K)$, the greater the incremental difference $\Pr(H/E \,\&\, K) - \Pr(E/K)$, which seems to validate the saying that the more surprising the evidence is, the more confirmational value it has. This observation, however, is double-edged, as we will see in chapter 5.

The problem of irrelevant conjunction, one of the main irritants of the HD method, is also illuminated. If $\{H, K\} \models E$, then also $\{H \,\&\, I, K\} \models E$, where I is anything you like, including a statement to which E is, intuitively speaking, irrelevant. But according to the HD account, E confirms $H \,\&\, I$. In a sense, the Bayesian analysis concurs, since if $\Pr(H \,\&\, I/K) > 0$, it

follows from the reasoning above that E incrementally confirms $H \,\&\, I$. However, note that it follows from (3.1) that the amounts of incremental confirmation that H and $H \,\&\, I$ receive are proportional to their prior probabilities:

$$\Pr(H/E \,\&\, K) - \Pr(H/K) = \Pr(H/K)[(1/\Pr(E/K)) - 1]$$

$$\Pr(H \,\&\, I/E \,\&\, K) - \Pr(H \,\&\, I/K) = \Pr(H \,\&\, I/K)[(1/\Pr(E/K)) - 1].$$

Since in general $\Pr(H \,\&\, I/K) < \Pr(H/K)$, adding the irrelevant conjunct I to H lowers the incremental confirmation afforded by E.

Finally, it is worth considering in a bit more detail the case of HD disconfirmation. Thus, suppose that when Nature speaks, she pronounces $\neg E$. If $\{H, K\} \models E$ and if K is held to be knowledge, then H must be false, so HD disconfirmation would seem to be equivalent to falsification. But as Duhem and Quine have reminded us, the deduction of observationally decidable consequences from high-level scientific hypotheses often requires the help of one or more auxiliary assumptions A. It is not fair to ignore this problem by sweeping the A's under the rug of K, since the A's are often every bit as questionable as H itself. Thus from Nature's pronouncement of $\neg E$ all that can be concluded from deductive logic alone is that $\neg H \vee \neg A$. If HD methodology were all there is to inductive reasoning, then there would be no principled way to parcel out the blame for the false prediction, and we would be well on the way to Duhem and Quine holism (see section 4 below). In particular, H could be maintained come what may if the only constraints operating were those that followed from direct observation and deductive logic. But the fact that the majority of scientists sometimes regard the maintenance of a hypothesis as reasonable and sometimes not is a fact of actual scientific practice that cries out for explanation. The Bayesian attempt at an explanation will be examined in section 7 below.

2 Hempel's Instance Confirmation

Having rejected the HD or prediction criterion of confirmation, Hempel constructed his own analysis of qualitative confirmation on a very different basis. He started with a number of conditions that he felt that any adequate theory of confirmation should satisfy, among which are the following:

Consequence condition If $E \models H$, then E confirms H.

Consistency condition If E confirms H and also H', then $\not\models \neg(H \,\&\, H')$.

Special consequence condition If E confirms H and $H \models H'$, then E confirms H'.

Hempel specifically rejected the converse consequence condition:

Converse consequence condition If E confirms H and $H' \models H$, then E confirms H'.

For to add the last condition to the first three would lead to the disaster that any E confirms any H.[3] (Note that HD confirmation satisfies the converse consequence condition but violates both the consistency condition and the special consequence condition.)

Hempel's basic idea for finding a definition of qualitative confirmation satisfying his adequacy conditions was that a hypothesis is confirmed by its positive instances. This seemingly simple and straightforward notion turns out to be notoriously difficult to pin down.[4] Hempel's own explication utilized the notion of the *development* of a hypothesis for a finite set I of individuals. Intuitively, $\mathrm{dev}_I(H)$ is what H asserts about a domain consisting of just the individuals in I. Formally, $\mathrm{dev}_I(H)$ for a quantified H is arrived at by peeling off universal quantifiers in favor of conjunctions over I and existential quantifiers in favor of disjunctions over I. Thus, for example, if $I = \{a, b\}$ and H is $(\forall x)(\exists y)Lxy$ (e.g., "Everybody loves somebody"), $\mathrm{dev}_I(H)$ is $(Laa \vee Lab) \,\&\, (Lbb \vee Lba)$. We are now in a position to state the main definitions that constitute Hempel's account.

Definition *E directly Hempel-confirms H* iff $E \models \mathrm{dev}_I(H)$, where I is the class of individuals mentioned in E.

Definition *E Hempel-confirms H* iff there is a class C of sentences such that $C \models H$ and E directly confirms each member of C.[5]

Definition *E Hempel-disconfirms H* iff E Hempel-confirms $\neg H$.

The difficulties with Hempel's account can be grouped into three categories. The first concerns the pillars on which the account was built: Hempel's so-called adequacy conditions. Bayesians have at least two ways of defining qualitative confirmation, one of which we already encountered in section 1; namely, *E incrementally confirms H* relative to K iff $\Pr(H/E \,\&\, K) >$

$\Pr(H/K)$. The second is an absolute rather than incremental notion; specifically, *E absolutely confirms H* relative to K iff $\Pr(H/E \,\&\, K) \geqslant k > .5$. (A third criterion sometimes used in the literature, e.g., Mackie 1963, says that E confirms H relative to K just in case $\Pr(E/H \,\&\, K) > \Pr(E/K)$. The reader can easily show that on the assumption that none of the probabilities involved is zero, this *likelihood criterion* is equivalent to the incremental criterion.) In both instances there appears to be a mismatch, since Hempel's account is concerned with a two-place relation 'E confirms H' rather than with a three-place relation ('E confirms H relative to K'). The Bayesians can accommodate themselves to Hempel either by taking K to be empty or by supposing that K has been learned and then working with the new probability function $\Pr'(\cdot) = \Pr(\cdot/K)$ obtained by conditionalization. But since one of the morals the Bayesians want to draw is that background knowledge can make a crucial difference to confirmation, I will continue to make K an explicit factor in the confirmation equation.

The first difficulty for Hempel's account can now be stated as a dilemma. For any choice of K compatible with H, Hempel's adequacy conditions accord well with the absolute notion of Bayesian confirmation. For example, if $\Pr(H/E \,\&\, K) > .5$ and $H \models H'$, then $\Pr(H'/E \,\&\, K) > .5$, so the special consequence condition is satisfied. But absolute confirmation cannot be what Hempel had in mind, since he holds that the observation of a single black raven a confirms the hypothesis that all ravens are black, even though for typical K's, $\Pr((\forall x)(Rx \to Bx)/Ra \,\&\, Ba \,\&\, K) \ll .5$. On the other hand, while the incremental concept of confirmation allows that a single instance can confirm a general hypothesis, both the consistency condition and the special consequence condition fail for not atypical K's, as examples by Carnap (1950) and Salmon (1975) show.[6] Of course, there may be some third probabilistic condition of confirmation that allows Hempel's account to pass between the horns of this dilemma. But it is up to the defender of Hempel's instance confirmation to produce the *tertium quid*. And even to conduct the search for a probabilistic *tertium quid* is to fall into the hands of the Bayesians.

The second category of difficulties revolves around the question of whether Hempel's account is too narrow. One reason for thinking so is that, as Hempel himself notes, a hypothesis of the form

$$(\forall x)(\exists y)Rxy \,\&\, (\forall x)(\forall y)(\forall z)[(Rxy \,\&\, Ryz) \to Rxz] \,\&\, (\forall x)\neg Rxx$$

cannot be Hempel-confirmed by any consistent E, since the development of such a hypothesis for a finite domain is inconsistent. Nor is the hypothesis $(\forall x)(\forall y)Rxy$ Hempel-confirmed by the set of evidence statements $\{Ra_i a_j\}$, where $i = 1, 2, \ldots, 10^9$ and $j = 1, 2, \ldots, 10^9 - 1$. Even more troublesome is the fact that Hempel's account is silent about how theoretical hypotheses are confirmed, for if, as Hempel intended, E is stated purely in the observational vocabulary and if H is stated in a theoretical vocabulary disjoint from the observational vocabulary, then E cannot, except in very uninteresting cases, Hempel-confirm H.[7] This silence is a high price to pay for overcoming some of the defects of the more vocal HD method.

Clark Glymour (1980) has sought to preserve Hempel's idea that hypotheses are confirmed by deducing positive instances of them from observation reports. In the case where H is stated in theoretical vocabulary, Glymour's bootstrapping method allows the deduction to proceed via auxiliary hypotheses, typically drawn from a theory T of which H itself is a part.[8] His basic confirmation relation is thus three-place: E confirms H relative to T.

The Bayesian response to these difficulties and to Glymour's reaction to them is twofold. First, there is no insuperable problem about how observational data can confirm, in either the incremental or absolute sense, a theoretical hypothesis; indeed, the application of Bayes's theorem shows just how such confirmation takes place, at least on the assumption that the prior probability of the hypothesis is nonzero (a matter that will be taken up in chapter 4). Second, unless bootstrap confirmation connects to reasons for believing the hypothesis or theory, it is of no interest. But once the connection is made, the bootstraps can be ignored in favor of the standard Bayesian account of reasons to believe. This matter will be examined in more detail in section 4 below.

The third category of difficulties is orthogonal to the second. Now the worry is that while Hempel's instance confirmation may be too narrow in some respects, it may be too liberal in other respects. Consider again the ravens hypothesis: $(\forall x)(Rx \rightarrow Bx)$. Which of the following evidence statements Hempel-confirm it?

E_1: Ra_1 & Ba_1

E_2: $\neg Ra_2$ & $\neg Ba_2$

E_3: $\neg Ra_3$

E_4: Ba_4

E_5: $\neg Ra_5$ & Ba_5

E_6: Ra_6 & $\neg Ba_6$

Only E_6 fails to Hempel-confirm the hypothesis, and that is because E_6 falsifies it. The indoor ornithology involved in using E_2 to E_5 as confirmation of the ravens hypothesis has struck many commentators as too easy to be correct. Bayesian treatments of Hempel's ravens paradox will be taken up in the following section.

If anything is safe in this area, it would seem to be that E_1 does confirm $(\forall x)(Rx \rightarrow Bx)$. But safe is not sure. Recall that Hempel's definition of confirmation is purely syntactical in that it is neutral to the intended interpretation of the predicates. This means that E_1 Hempel-confirms $(\forall x)(Rx \rightarrow Bx)$ even if we take Bx to mean not that x is black but that x is blite, i.e., x is first examined before the year 2000 and is black, or else is not examined before 2000 and is white. Let a_i be first examined in the year i. Then by the special consequence condition, Ra_1 & Ba_1 & Ra_2 & Ba_2 & … & Ra_{1999} & Ba_{1999} Hempel-confirms the prediction $Ra_{2001} \rightarrow Ba_{2001}$, i.e., the prediction that if a_{2001} is a raven, then it is white, which is, to say the least, counterintuitive. We have here an instance of what Goodman (1983) calls the "new riddle of induction." The Bayesian treatment of this problem will be given in detail in chapter 4. But for now I will simply note on behalf of the Bayesians that they are not committed to assigning probabilities purely on the basis of the syntax of the hypothesis and the evidence, as Hempel's analogy between deductive and inductive logic would suggest. The present example is enough to show that an adequate account of confirmation must be sensitive to semantics, and this lesson is easily incorporated into Bayesianism.

3 The Ravens Paradox

In sections 1 and 2 Bayesianism gained reflected glory of sorts from the whippings the HD and Hempel accounts took. It is time for Bayesianism to earn additional glory of a more positive sort.

Hempel took it as a desirable consequence of his account that the evidence Ra & Ba confirms the hypothesis $(\forall x)(Rx \rightarrow Bx)$.[9] The paradox of the ravens in one of its forms arises from the fact that on Hempel's analysis, the evidence $\neg Rb$ & $\neg Bb$ also confirms $(\forall x)(Rx \rightarrow Bx)$. Before turning to the Bayesian analysis of the paradox itself, it is worth noting that the Bayesian is not even willing to go the first step with Hempel without first looking both ways.

Suppose that $0 < \Pr(H/K) < 1$, where H stands for the ravens hypothesis. Then by an application of Bayes's theorem it follows that finding a to be a black raven induces incremental confirmation,

$$\Pr(H/Ra \,\&\, Ba \,\&\, K) > \Pr(H/K),$$

just in case

$$\Pr(Ra/H \,\&\, K) > \Pr(Ra/\neg H \,\&\, K) \times \Pr(Ba/Ra \,\&\, \neg H \,\&\, K).$$

Incremental disconfirmation results just in case the inequality is reversed.[10] The reader is invited to reflect on the kinds of background knowledge K that will make or break these inequalities. Consider, for instance, a version of I. J. Good's (1967) example. We are supposed to know in advance (K) that we belong to one of two bird universes: one where there are 100 black ravens, no nonblack ravens, and 1 million other birds, or else one where there are 1,000 black ravens, 1 white raven, and 1 million other birds. Bird a is selected at random from all the birds and found to be a black raven. This evidence, Good claims, undermines the ravens hypothesis. Use the above formula to test this claim. Such exercises help to drive home the point that a two-place confirmation relation that ignores background evidence is not very useful.

Let us turn now to the Bayesian treatment of the bearing of the evidence of nonblack nonravens on the ravens hypothesis. Suppes (1966) invites us to consider an object a drawn at random from the universe. Set

$$\begin{aligned} &\Pr(Ra \,\&\, Ba/K) = p_1, \qquad \Pr(Ra \,\&\, \neg Ba/K) = p_2, \\ &\Pr(\neg Ra \,\&\, Ba/K) = p_3, \qquad \Pr(\neg Ra \,\&\, \neg Ba/K) = p_4. \end{aligned} \tag{3.2}$$

Then

$$\Pr(\neg Ba/Ra \,\&\, K) = p_2/(p_1 + p_2) \tag{3.3}$$

and

$$\Pr(Ra/\neg Ba \,\&\, K) = p_2/(p_2 + p_4). \tag{3.4}$$

From (3.3) and (3.4) it follows that $\Pr(\neg Ba/Ra \,\&\, K) > \Pr(Ra/\neg Ba \,\&\, K)$ iff $p_4 > p_1$. But from what we know of the makeup of our universe, it seems safe to assume that $p_4 \gg p_1$, with the consequence that the conditional probability of a's being nonblack, given that it is a raven, is much greater than the conditional probability of a's being a raven, given that it is nonblack. The moral Suppes wants us to draw from this is that sampling from the class of ravens is more productive than sampling from the class of nonblack objects, since the former procedure is more likely to produce a counterexample to the ravens hypothesis.

There are two qualms about this moral. The first is that it doesn't seem directly useful to Bayesians; indeed, at first blush it seems more congenial to a Popperian line that emphasizes the virtues of attempted falsifications of hypotheses. Second, it is not clear how the moral follows from the inequality derived, since a was supposed to result from a random sample of the universe at large rather than from a random sample of either the class of ravens or the class of nonblack objects.

Horwich's (1982) attack on the ravens paradox starts from the observation that there are several ways to obtain the evidence $Ra \,\&\, Ba$, namely, to pick an object at random from the universe at large and find that it has both ravenhood and blackness, to pick an object at random from the class of ravens and find that it is black, or to pick an object at random from the class of black things and find that it is a raven. A similar remark applies to the evidence $\neg Rb \,\&\, \neg Bb$. Horwich introduces the notation R^*a to mean that a was drawn at random from the class of ravens and the notation $\neg B^*b$ to mean that b was drawn at random from the class of nonblack things. To illuminate the ravens paradox, he wants to compare the confirmational effects of the two pieces of evidence $R^*a \,\&\, Ba$ and $\neg B^*b \,\&\, \neg Rb$. According to Horwich's application of Bayes's theorem,

$$\Pr(H/R^*a \,\&\, Ba \,\&\, K) = \Pr(H/K)/\Pr(R^*a \,\&\, Ba/K) \tag{3.5}$$

and

$$\Pr(H/\neg B^*b \,\&\, \neg Rb \,\&\, K) = \Pr(H/K)/\Pr(\neg B^*b \,\&\, \neg Rb/K), \tag{3.6}$$

where K is the same as before. Thus

$$\begin{aligned} &\Pr(H/R^*a \,\&\, Ba \,\&\, K) > \Pr(H/\neg B^*b \,\&\, \neg Rb \,\&\, K) \\ &\quad \text{iff} \quad \Pr(\neg B^*b \,\&\, \neg Rb/K) > \Pr(R^*a \,\&\, Ba/K). \end{aligned}$$

But the latter is true for our universe, Horwich asserts.

But as with Suppes's construction, it is not clear how this conclusion follows. In the first place, why is it true (as (3.5) and (3.6) assume) that

$$\Pr(R^*a \;\&\; Ba/H \;\&\; K) = \Pr(\neg B^*b \;\&\; \neg Rb/H \;\&\; K) = 1?$$

It is true that the probability of a randomly chosen raven being black, given H & K, is 1, but $\Pr(R^*a \;\&\; Ba/H \;\&\; K)$ is the probability that an object a is randomly chosen from the class of ravens and is black, given H & K, and this probability is surely not 1. In the second place, comparing $\Pr(\neg B^*\mathrm{b} \;\&\; \neg Rb/K)$ and $\Pr(R^*a \;\&\; Ba/K)$ involves a comparison of the probability that an object will be randomly sampled from the class of ravens with the probability that it will be randomly sampled from the class of nonblack things, and such a comparison seems peripheral to the paradox at best.

Horwich's basic idea can be brought to fruition by putting into the background knowledge $\hat{K}$ the information that R^*a and $\neg B^*b$. Bayes's theorem can then be legitimately applied to the new $\hat{K}$ to conclude that

$$\Pr(H/Ra \;\&\; Ba \;\&\; \hat{K}) = \Pr(H/\hat{K})/\Pr(Ba/\hat{K}) \tag{3.7}$$

and

$$\Pr(H/\neg Rb \;\&\; \neg Rb \;\&\; \hat{K}) = \Pr(H/\hat{K})/\Pr(\neg Rb/\hat{K}). \tag{3.8}$$

Thus, relative to this $\hat{K}$, the evidence Ra & Ba has more confirmational value vis-à-vis the ravens hypothesis than does $\neg Rb$ & $\neg Bb$ just in case $\Pr(\neg Rb/\hat{K}) > \Pr(Ba/\hat{K})$. A further application of the principle of total probability shows that this latter inequality holds just in case $\Pr(\neg Ba/\neg H \;\&\; \hat{K}) > \Pr(Rb/\neg H \;\&\; \hat{K})$. This last inequality presumably does hold in our universe, for given that some ravens are nonblack ($\neg H$), we are more likely to produce one of them by sampling from the class of ravens than by sampling from the class of nonblack things simply because of the known size and heterogeneity of the class of nonblack things as compared with the known size of the class of ravens. Suppes is thus vindicated after all, since the greater confirmatory power of Ra & Ba over $\neg Rb$ & $\neg Bb$ has to do with the relative threats of falsification. In this way Bayesianism pays a backhanded compliment to Popper's methodology; namely, it is precisely because, contrary to Popper, inductivism is possible that the virtues of sincere attempts to falsify can be recognized.[11]

Similar points are made by Gaifman (1979), although his assumed sampling procedure is somewhat different. Let $\tilde{K}$ report that c was drawn at random from the universe and found to be a raven and that d was also drawn at random from the universe and found to be nonblack. An analysis like the one above shows that

$$\Pr(H/Rc \;\&\; Bc \;\&\; \tilde{K}) > \Pr(H/\neg Rd \;\&\; \neg Bd \;\&\; \tilde{K})$$

just in case $\Pr(\neg Bc/\neg H \;\&\; \tilde{K}) > \Pr(Rd/\neg H \;\&\; \tilde{K})$.

But the procedure of sampling from the universe at large can be wasteful, since it can produce relatively useless results, such as $\neg Re$ & Be. Moreover, one can wonder whether the evidence Ra & Ba, under the assumption that a was drawn at random from the class of ravens, gives better confirmational value than the evidence Rc & Bc, under the assumption that c was drawn at random from the universe at large, i.e., whether

$$\Pr(H/Ra \;\&\; Ba \;\&\; \hat{K} \;\&\; \tilde{K}) > \Pr(H/Rc \;\&\; Bc \;\&\; \hat{K} \;\&\; \tilde{K}).$$

I leave it to the reader to ponder this question with the clue that the answer is positive just in case

$$\Pr(\neg Ba/\neg H \;\&\; \hat{K} \;\&\; \tilde{K}) > \Pr(\neg Bc/\neg H \;\&\; \hat{K} \;\&\; \tilde{K}).$$[12]

4 Bootstrapping and Relevance Relations

In *Theory and Evidence* (1980) Glymour saw bootstrapping relations not only as a means of extending Hempel's instance confirmation to theoretical hypotheses but also as an antidote to Duhem and Quine holism. It makes a nice sound when it rolls off the tongue to say that our claims about the physical world face the tribunal of experience not individually but only as a corporate body. But scientists, no less than business executives, do not typically act as if they are at a loss as to how to distribute praise through the corporate body when the tribunal says yea, or blame when the tribunal says nay. This is not to say that there is always a single correct way to make the distribution, but it is to say that in many cases there are firm intuitions. Bootstrap relations would help to explain these intuitions if they helped to explain why it is that for some but not all H's that are part of a theory T, E bootstrap-confirms H relative to T.

As a sometime Bayesian I now think that bootstrapping should be abandoned in favor of a Bayesian analysis. Bayesians can be sympathetic to the two motivations for bootstrapping mentioned above in section 2. At the same time Bayesians can recognize that any account of confirmation modeled on Hempel's approach will have two fatal flaws. (1) For Hempel, whether or not E confirms H depends only on the syntax of E and H. But from Goodman we know this to be wrong (see section 2 above and chapter 4). (2) For Hempel, confirmation is a two-place relation. But from the ravens paradox and other examples we know that background information K must be brought into the analysis to get an illuminating treatment. The relevance of these points to bootstrapping can be brought into focus with the help of Christensen's (1983) examples.

Let T have as its axioms H_1: $(\forall x)(Rx \rightarrow Bx)$, and H_2: $(\forall x)(Rx \rightarrow Hx)$, the former of which is our old friend the ravens hypothesis and the latter of which asserts that all ravens live a happy afterlife in bird heaven. At first blush, evidence from the observation of the color of a raven is directly relevant to H_1 but is irrelevant to H_2, even relative to T. But Christensen shows how, with a little logical flimflam, such evidence leads to a bootstrap confirmation of H_2 relative to T. On the standard conception of theories, T is the logical closure of $\{H_1, H_2\}$. Thus it is part of T that

H_3: $(\forall x)[Rx \rightarrow (Bx \leftrightarrow Hx)]$.

From E: Ra & Ba, we can deduce via H_3 that Ra & Ha, which is a Hempel positive instance of H_2. Moreover, the possible alternative evidence E': Ra & $\neg Ba$, leads via H_3 to Ra & $\neg Ha$, which is a counterinstance of H_2. Together these "computations" constitute a positive bootstrap test of H_2. But intuitively, only phony-baloney confirmation/testing has taken place. A revised set of bootstrap conditions proposed by Glymour (1983) rule out this particular example, but Christensen (1990) has shown how the counterexamples can be revived in a more complicated form.[13]

One could seek further restrictions to rule out the new counterexamples, but this now seems to me to be a mistake—we do not want a once-and-for-all answer to, Does E confirm H relative to T? that is independent of the interpretation of the nonlogical constants in E, H, and T and also independent of the background knowledge.

To make the point more concrete, let me use another of Christensen's examples, which is structurally identical to the above ravens case. Now T is the logical closure of H_1: $(\forall x)(Sx \rightarrow Ax)$, and H_2: $(\forall x)(Sx \rightarrow Vx)$. H_1 is intended to assert that anyone with certain disease symptoms has the antibodies to a certain virus, while H_2 is intended to assert that anyone with the said symptoms has been infected by the said virus. T contains

H_3: $(\forall x)[Sx \rightarrow (Ax \leftrightarrow Vx)]$. The evidence E: Sa & Aa, leads via H_3 to a positive instance of H_2, while the alternative possible evidence E': Sa & $\neg Aa$, leads via H_3 to a negative instance of H_2. Although structurally identical to the former example, we are not so ready to see phony-baloney confirmation/testing here.

To diagnose the felt asymmetries between the two cases, we need to know to what end the three-place Glymourian relation 'E bootstrap-confirms/tests H relative to T' is to be put. E bootstrap-confirms H relative to T cannot be taken to imply that, assuming T to be true or well confirmed, E confirms H, for in the cases at issue H is part of T. Rather, the most plausible usage is in adjudicating questions of evidential relevance. Note that in these examples Hempel's version of the "prediction criterion" of confirmation is satisfied; i.e., E is of the form E_1 & E_2, where $\{T, E_1\} \models E_2$ but $E_1 \not\models E_2$, while E' is of the form E'_1 & E'_2, where $\{T, E_1\} \models \neg E'_2$ but $E'_1 \not\models \neg E_2$. The antiholist then asks, if E is found to hold, to which parts of T can the praise for the successful prediction be attributed? If E' is found to hold, on which parts of T can the blame for the unsuccessful prediction be laid?

With this interpretation of bootstrapping, the Bayesian diagnosis of the counterexamples is straightforward. H gets praise from E if, relative to K, E incrementally confirms H, and H gets blame from E' if, relative to K, H is incrementally disconfirmed by E'. The bird-heaven case gave off a bad odor since our current background knowledge K would have to be radically altered for Ra & Ba to incrementally confirm H_2 or for Ra & $\neg Ba$ to incrementally disconfirm H_2. Indeed, given the tenets of traditional empiricism, we could never get to an alternative K where this would happen. By contrast, the virus case smelled sweeter even though, from the point of view of bootstrapping, it is structurally identical to the bird-heaven case. A possible reason is that in the virus case H_2 will get praise from Sa & Aa and blame from Sa & $\neg Aa$ if K makes likely the proposition that all and only those people who have been infected by the virus have antibodies to it, a not implausible situation.

It might be complained that while such a diagnosis does in fact help to explain intuitions, it is irrelevant to the original project; the aim of that project was to provide an internalist analysis of relevance relations, and given that aim, it is illegitimate to bring in K. The response to this complaint parallels the response to Hempel's complaint that background information about the relative sizes of the classes of ravens and nonblack things is irrelevant to his project, which concerns only the two-place relation '*E* confirms *H*'; namely, no interesting account of confirmation can be developed if *K* is left out of the picture.

Aron Edidin (1988) has maintained that the core of the program of relative confirmation is left untouched by Christensen's examples. I think that there is a sense in which Edidin's contention is correct, but by the same token I think that the program of relative confirmation can be seen to be drained of much of its interest. Let us suppose that the core of the program is concerned with the relation '*E* confirms *H* relative to auxiliaries *A*', where typically the auxiliaries do not include *H* itself. Edidin's point is that there is nothing in Christensen's examples to suggest that the apparatus developed in *Theory and Evidence* is not adequate to provide a correct explication of this relation. Thus in Christensen's ravens example there is nothing counterintuitive to maintaining that E: Ra & Ba, does confirm H_2: $(\forall x)(Rx \rightarrow Hx)$, relative to the auxiliary assumption H_3: $(\forall x)[Rx \rightarrow (Bx \leftrightarrow Hx)]$. This seems to me correct in the following respect: in the sense in which Hempel could say that E: Ra & Ba, confirms H_1: $(\forall x)(Rx \rightarrow Bx)$, it is also natural by extension to say that E confirms H_2 relative to H_3.

But if the core of the program of relative confirmation is left untouched, it remains to ask what purpose is served by the program. Two responses suggest themselves. First, we can hope to use the relation '*E* confirms *H* relative to *A*' to explicate theory-relative confirmation. Thus, we can say that '*E* confirms *H* relative to *T*', where *T* typically contains *H*, means that there is an appropriate *A* in *T* such that *E* confirms *H* relative to *A*. Here the appropriateness of *A* is supposed to guarantee that the resulting confirmation/disconfirmation of *H* relative to *T* by *E* implies that the praise/blame for *T*'s passing/failing to pass an HD test can be attached to *H*. The presumption of *Theory and Evidence* was that the appropriateness of *A* can be settled purely in terms of structural relations among *A*, *H*, *E*, and *T*. This presumption is belied by the analysis above of Christensen's examples, which shows that the parceling out of praise and blame depends on the epistemic status of *A*, which in turn depends upon the background knowledge.

The second response is that getting a handle on relative confirmation is useful in deciding how evidence affects the credibility of hypotheses and in turn the credibility of theories of which the hypotheses are parts. But again, the epistemic status of the auxiliaries must be taken into account. Edidin's discussion indicates that the move from '*E* confirms *H* relative to *A*' to '*E* contributes to the credibility of *H*' is a tricky one; it requires not only that the auxiliaries *A* "must themselves be credible." In some cases it requires also that "their credibility must be substantially independent of the credibility of the evidence" (p. 268) and in other cases that they have "antecedent credibility independent of that of the hypotheses" (p. 269). But what exactly do these requirements come to? I submit that no precise answer can be given without invoking the Bayesian apparatus. Further, the answer this apparatus yields is that no answer can be given in the abstract: it depends on the background information *K*, and it depends not just on the logico-structural relations involved in the HD and bootstrapping account of relative confirmation but also on the intended interpretation of the nonlogical terms in *E*, *H*, and *A*.

The complaint here is not that, on pain of circularity, HD or bootstrapping relations of relative confirmation cannot figure in an account of how evidence bears on the credibility of theoretical hypotheses; rather, the complaint is that such relations may not contribute in any perspicuous way to the assessment of that bearing. Consider again the simpler case of the confirmation of observational hypotheses. How, for example, does evidence about the color of ravens and nonravens bear on the credibility of the hypothesis that all ravens are black? By now, I hope, the reader is convinced that an illuminating path to an answer need not take the form of first deciding when *E* Hempel-confirms *H* and then trying to puzzle out the further conditions necessary for the move from Hempel-confirmation to an incremental increase in credibility. The moral here has double strength when we move from Hempel-confirmation of observational hypotheses to the more complicated case of relative confirmation of theoretical hypotheses.

5 Variety of Evidence and the Limited Variety of Nature

It is a truism of scientific methodology that variety of evidence can be as important or even more important than sheer amount of evidence. An adequate account of confirmation is not under obligation to give an unqualified endorsement to all such truisms, but it should be able to identify the valid rationale (if any) of such truisms.

A Bayesian explanation of the virtue of variety of evidence would concentrate on the ability of variety to contribute to a significant boost in the posterior probability of a hypothesis. To illustrate how part of the explanation might go, consider again the HD case where $H, K \models E$, and suppose that E is E_1 & E_2 & ... & E_n, where the E_i report the outcomes of performing some one experiment over and over or alternatively the outcomes of a series of different experiments. The most helpful form of Bayes's theorem to cover this situation is

$$\Pr(H/E \,\&\, K) = \frac{\Pr(H/K)}{\Pr(E_1/K) \times \Pr(E_2/E_1 \,\&\, K) \times \cdots \times \Pr(E_n/E_1 \,\&\, \ldots \,\&\, E_{n-1} \,\&\, K)}. \tag{3.9}$$

As we will see in chapter 4, if $\Pr(H/K) > 0$, the factor $\Pr(E_n/E_1 \,\&\, \ldots \,\&\, E_{n-1} \,\&\, K)$ must go to 1 as n grows without bound. This factor gives the probability of the next experimental outcome predicted by H, conditional on the background information K and the information that the previous predictions have been borne out. The more slowly this probability approaches 1, the smaller the denominator (for a given n) and hence the larger the posterior probability of H (for a given n). This is exactly where variety of evidence enters, for the more various the experiments, the slower one would expect the approach to certainty to be for the next outcome.[14] At one extreme is the case where the E_i are the outcomes of repeating the same experiment consisting, say, of measuring over and over again a quantity believed to have a stable value. Then with appropriate assumptions K about the reliability of the measuring apparatus, only a few repetitions are needed to achieve near certainty for the next instance, and amassing a large number of further instances achieves little gain for the posterior probability of H. At the other extreme is the case where the E_i are the outcomes of experiments that are not only different but seem quite unrelated. Then new instances will make for a bigger gain in the posterior probability of H.[15]

These remarks have value only if we already have a grip on the notion of variety of evidence. But rather than trying to give an independent analysis of variety, what I would like to suggest is that the observations above can be given a new twist and used to define 'variety of evidence' through rate of increase in the factors

$\Pr(E_n/E_1 \,\&\, \ldots \,\&\, E_{n-1} \,\&\, K)$.[16]

Such an analysis has two consequences, one of which is obvious, the other of which is a little surprising.

The obvious consequence is that the notion of variety of evidence has to be relativized to the background assumptions K, but there is no more than good scientific common sense here, since, for example, before the scientific revolution the motions of the celestial bodies seemed to belong to a different variety than the motions of terrestrial projectiles, whereas after Newton they seem like peas in a pod.

The less obvious consequence is that induction, or a necessary condition for it, presupposes a limited variety in nature, as Keynes (1962) tried to teach us. As already remarked, $\Pr(H/K) > 0$, which is necessary for the probabilification of H, implies that

$$\Pr(E_n/E_1 \,\&\, \ldots \,\&\, E_{n-1} \,\&\, K) \rightarrow 1$$

as $n \rightarrow \infty$. This means that from the point of view of the proposed analysis of variety, E_n for large enough n cannot be counted as various with respect to $E_1, E_2, \ldots, E_{n-1}$, contrary to what our untutored intuitions might have told us. The fact that the E_i are unified in the very minimal sense of being entailed by a single H to which we assign a nonzero prior eventually forces us to see them as nonvarious.

Another aspect of the importance of variety of evidence arises in conjunction with eliminative induction, whose virtues are touted in chapter 7. Bayes's theorem in the form (2.2) shows how the probability of a hypothesis is boosted by evidence that eliminates rival hypotheses. Thus variety of evidence can be analyzed from the point of view of how likely the evidence is to produce efficient elimination.[17]

6 Putnam and Hempel on the Indispensability of Theories

Induction by enumeration is inadequate for capturing many of the inferences routinely made in the advanced sciences, as is brought out very nicely by the following example of Putnam's (1963a). Imagine that you were a member of the Los Alamos Project during World War II. As you prepare for the first test of what you hope will be an atomic bomb, you consider prediction H: when these two subcritical masses of U_{235} are slammed together to form a supercritical mass, there will be an atomic explosion. H has a counterpart in purely observational terms, namely H': when these two rocks are slammed together, there will be a big bang. If E is the sum of the directly relevant observations made up to this juncture, there is no way for an inductivist who limits himself to simple enumeration to move from E to a confidence in H'. For up to now there have been no recorded cases of rocks of this kind exploding, but there have been many recorded cases of rocks of this kind being slammed together without exploding (because critical mass was never reached). Nevertheless, you and your fellow project scientists are confident of H'. Why?

The Bayesian is happy to supply the answer. You were in possession of a theory T of the atomic nucleus that entails H'. Applying the principle of total probability to the total available observational evidence E & $\hat{E}$ gives

$$\Pr(H'/E \,\&\, \hat{E}) = \Pr(T/E \,\&\, \hat{E}) + \Pr(H'/\neg T \,\&\, E \,\&\, \hat{E}) \times \Pr(\neg T/E \,\&\, \hat{E}).$$

Thus if your opinions conformed to the probability calculus, your confidence in H' should have been at least as great as your confidence in T. And the combination of E and $\hat{E}$ made you somewhat confident of T (because, for example, T entails other experimental regularities whose positive instances are recorded by $\hat{E}$). Further, $\neg T$ includes other theories that also entail H' or make H' highly probable, and E & $\hat{E}$ made you somewhat confident of those theories. The upshot was that you were more than somewhat confident of H'.

Putnam used this story to register a complaint against any explication of degree of confirmation that makes the confirmation of H' on E & $\hat{E}$ independent of the presence or absence in the language of predicates not occurring in H', E, or $\hat{E}$ (what Carnap in 1950 and 1952 called an inductive method of the "first kind"). In terms of the present example, such an explication implies that $\Pr(H'/E \,\&\, \hat{E})$ can be assessed in a language that contains only observational predicates. But since expressions involving T cannot occur in such a language, the explanation above of the expectations of the Los Alamos scientists cannot be stated in such a setting. To provide an explanation within the strictures of an inductive method of the first kind, it must be supposed that the scientists involved would have had the same degree of confidence in H' had they never considered T, a highly implausible supposition, to say the least. Of course, it could be replied that the failure of inductive methods of the first kind to accord with the actual psychology of scientists may be ignored, since the task of explicating degree of confirmation is a normative rather than a descriptive one. The rejoinder is that the normative status of a proposed explication comes into question when the explication fails to accord with what the history of science provides as paradigm cases of good inferences. In effect, Carnap agreed with this rejoinder in his response to Putnam. He wrote that

> for situations of this kind we must construct a new inductive logic which refers to the theoretical language instead of the observational language. I would say that the scientists at the time in question would indeed have been willing to bet on the positive success of the first nuclear explosion on the basis of the available evidence, including results of the relevant laboratory experiments. Inductive logic must reconstruct this willingness by ascribing to $c(H, E)$ a considerable positive value.[18] (1963b, p. 988)

Is there an argument here for scientific realism? Not much of one, but something is better than nothing. Consider the position of an antirealist who is neither an instrumentalist nor an inductive skeptic with respect to observational predictions but who is an inductive skeptic with respect to theoretical claims. In the Los Alamos example such an antirealist will agree that reasonable expectations about the explosion prediction H' can be formed on the basis of E & $\hat{E}$. He also agrees that the nuclear theory T has a truth value and that the proposition asserting that T is true is not merely a disguised way of asserting that observational predictions of T are correct. But he nevertheless denies that the observational evidence E & $\hat{E}$ serves as a basis for a reasonable belief in the truth of T. Such an antirealist is very much in the same position as someone who uses a Carnapian method of the first kind, and whatever objections can be brought against the latter can also be brought against the former.

The above considerations also help to illuminate Hempel's (1958) proposed resolution of the "theoretician's dilemma." On Hempel's formulation, the dilemma runs thus: either theoretical terms fulfill their function of systematizing deductive connections among observation statements or they don't. If they don't, they are obviously dispensable. If they do, they are likewise dispensable, since Craig's (1956) lemma shows that the observational consequences of an axiomatizable theory can always be reaxiomatized in purely observational vocabulary. Hence theoretical terms are dispensable. Hempel's response was that theories may be indispensable because they serve to establish *inductive* as well as *deductive* connections.

T might be said to be essential to establishing inductive connections among observables if there are observation sentences O_1 and O_2 such that $\Pr(O_2/T \,\&\, O_1) > \Pr(O_2/O_1)$, or more interestingly, if $\Pr(O_2/T \,\&\, O_1) > \Pr(O_2/O_T \,\&\, O_1)$, where O_T is a sentence logically equivalent to the set of observational consequences of T.[19] The first condition is certainly satisfied in the Los Alamos example with $O_1 = E \,\&\, \hat{E}$ and $O_2 = H'$, and for sake of argument we may suppose that the second condition is satisfied as well.

But on further reflection, these facts do not by themselves establish the claimed indispensability of T. In the Los Alamos example, the key question is what degree of confidence to put in H' on the basis of the total available evidence E & $\hat{E}$. Thus in this example the claim that theories are indispensable for purposes of inductive systematization must be understood as the claim that the evaluation of $\Pr(H'/E \,\&\, \hat{E})$ depends in some essential way on T. But what way is this? I suggest that the answer must be the one supplied by my discussion of Putnam's story. And I would further suggest that the moral of the story can be generalized.

Suppose that for purposes of scientific investigation of a certain domain, an inductive agent adopts a language $\mathscr{L}$ and a degree-of-belief function Pr on the propositions $\mathscr{A}$ of $\mathscr{L}$. We may suppose that $\mathscr{L}$ is a purely observational language. Subsequently the agent expands her language to $\mathscr{L}'$, which includes theoretical predicates, and adopts a degree-of-belief function Pr′ for the propositions $\mathscr{A}' \supset \mathscr{A}$ of the new language. Even though she is a rational agent, it may very well be that Pr′ restricted to $\mathscr{A}$ does not coincide with her previous belief function Pr. Of course, this phenomenon has nothing to do per se with the observational/theoretical distinction; it is merely a corollary of the point that the probability assigned to a proposition may depend upon the possibility set in which the proposition is imbedded. The moral here has an intralanguage counterpart. Within, say, the language of physics as it is constituted at any particular time, physicists are explicitly aware of only a small portion of the possible theories that can be formulated in the language. When new theories are formulated, the range of the explicitly recognized possibilities being thereby expanded, the probabilities of previously considered hypotheses and theories may change. This matter is taken up in chapters 5 and 7.

A striking consequence emerges when we combine such morals with Carnap's principle of tolerance, according to which "everyone is free to use the language most suited to his purpose" (1963a, p. 18). Since the exercise of this freedom is guided to a large extent by pragmatic factors, and since degree of confirmation is affected by the choice of language, the implication is that evidential support has a pragmatic dimension. Pure personalists will hardly be shocked by this consequence, but those who want confirmation theory to deliver rational and objective degrees of belief may not be so shock-proof.

Those who do find such a consequence repugnant may want to consider restrictions on the principle of tolerance, but it is hard to see how a principled intolerance is to be implemented. Alternatively, the consequence can be avoided by doing confirmation theory in a universal language adequate for reconstructing all past and future scientific endeavors. But even if such a utopian scheme is possible, its relevance to the actual practice of science, which takes place in a context far from utopia, is tenuous.[20] Rather than try to avoid the consequence, I recommend a cautious embrace. Chapter 7 gives a concrete example of one form the embrace might take.

7 The Quine and Duhem Problem

If hypotheticodeductivism were the only tool available for assessing evidence, we would be at a loss in making judgments about how evidence bears differentially on the components of a scientific theory. Some additional tool is thus sorely needed. In section 4, I found fault with Glymour's attempt to parcel out praise and blame using bootstrapping relations, and I intimated that the parceling out is best accomplished with Bayesian means. Sometimes a Bayesian analysis supports a kind of holism. Thus if T consists of the conjunction of T_1 and T_2, and if T contradicts E & K, the blame may attach to T as a whole without sticking to either component T_1 or T_2. Indeed, Wesley Salmon (1973) has provided an example where, relative to K, E incrementally confirms each of T_1 and T_2, even though T is refuted by E & K.[21] In more typical cases of refutation, however, our intuitions suggest that the blame does stick to one or another component of the theory and also that it sticks more firmly to some components than to others.

An example of how the Bayesian apparatus can be used to support such intuitions in historically realistic cases has been given by Jon Dorling (1979). Suppose that theory T consists of core hypotheses T_1 and auxiliary assumptions T_2; that $T_1 \,\&\, T_2 \models E'$; and finally that nature pronounces E, which is incompatible with E'.[22] Dorling assumes that T_1 is probabilistically irrelevant to T_2 (that is, $\Pr(T_2/T_1) = \Pr(T_2)$), that the priors $\Pr(T_1) = k_1$ and $\Pr(T_2) = k_2$ satisfy $k_1 > k_2$ and $k_1 > .5$, while the likelihoods $\Pr(E/\neg T_1 \,\&\, T_2) = k_3$, $\Pr(E/T_1 \,\&\, \neg T_2) = k_4$, and $\Pr(E/\neg T_1 \,\&\, \neg T_2) = k_5$ satisfy $k_3 \ll k_4, k_5 \ll 1$. Then Bayes's theorem shows that the blame falls more heavily on the auxiliaries T_2 than on the core T_1. If we take the time to be the mid nineteenth century, T_1 to be Newton's theory of motion and gravitation, T_2 the assumption that tidal effects do not influence lunar secular acceleration, and E the observed secular acceleration of the moon, then Dorling argues that plausible values of the relevant probabilities are $k_1 = .9$, $k_2 = .6$, $k_3 = .001$, $k_4 = k_5 = .05$. With these values he finds that $\Pr(T_1/E) = .8976$ and $\Pr(T_2/E) = .003$, so that the refuting evidence E only slightly reduces the probability of the core of the theory, while strongly undermining the auxiliary.[23]

Assuming that Dorling's reconstruction of the prevailing degrees of belief is historically correct, we are presented with a Bayesian success story in the form of an explanation of the attitudes and behavior displayed by the scientific community during an important incident in nineteenth-century astronomy. But what we don't yet have is a solution to the Quine and Duhem problem, at least not if what we demand of a solution is a demonstration that one way of parceling out the blame is rationally justified while others are not. For it is perfectly compatible with Bayesian personalism to assign values to k_1 through k_5 that make T_1 the goat while rendering T_2 blameless.[24] We have arrived at one aspect of the general problem of the objectivity of scientific inference, a problem that will occupy us from chapter 6 onward. I will note in advance that while much of the attention on the Bayesian version of this problem has focused on the assignments of prior probabilities, the assignments of likelihoods involves equally daunting difficulties.

In the present context the difficulties can be illustrated by noting that when $T_1 \,\&\, T_2 \vdash \neg E$ but nature pronounces E, then blame attaches squarely to T_1 in the sense that $\Pr(T_1/E) \ll \Pr(T_1)$ just in case

$$\Pr(E/T_1 \,\&\, \neg T_2) \times \Pr(\neg T_2/T_1) \ll \Pr(E/\neg T_1 \,\&\, T_2) \times \Pr(T_2/\neg T_1) + \Pr(E/\neg T_1 \,\&\, \neg T_2) \times \Pr(\neg T_2/\neg T_1).$$

In general, none of the factors involved has an objective character, and a large variability can be expected in the values assigned by different persons. Dorling's argument that this inequality fails in his historical case study is based on the assumption that $\Pr(E/\neg T_1 \,\&\, T_2)$ is small—an assumption Dorling takes to be justified because (he says) no plausible rival to Newton's theory could predict E either quantitatively or qualitatively. This justification succeeds if $\neg T_1$ is limited to rivals actually constructed by nineteenth-century physicists. But a critic of this analysis might well ask why pronouncements about what it is and isn't rational to believe in the

face of E should depend on the vicissitudes of which of the myriad possible theories happened to be constructed by physicists of the time.

Let us attempt to add some objectivity by moving to a simple if unrealistic case. Assume first that T_1 and T_2 are probabilistically irrelevant to one another. Assume second that we can parse $\neg T_1$ as $T_1^1 \vee T_1^2 \vee \ldots \vee T_1^n$, where the T_1^i are pairwise inconsistent, and that we can parse $\neg T_2$ as $T_2^1 \vee T_2^2 \vee \ldots \vee T_2^m$, where the T_2^j are also pairwise inconsistent. Assume finally that T_1 or any one of the T_1^i when conjoined with T_2 or any one of the T_2^k together entail a definite prediction for the phenomenon in question. Then the condition for blame to attach to T_1 becomes

$$\sum_j \Pr(T_2^j) \ll \left[\frac{\Pr(T_2)}{\Pr(\neg T_1)}\right] \times \sum_i \Pr(T_1^i) + \left[\frac{\Pr(\neg T_2)}{\Pr(\neg T_1)}\right] \times \sum_k \Pr(T_1^k),$$

where the sum on j is taken over values such that $T_1 \,\&\, T_2^j \vdash E$, the sum on i is taken over values such that $T_1^i \,\&\, T_2 \vdash E$, and the sum on k is taken over values such that $T_1^k \,\&\, \neg T_2 \vdash E$ (i.e., $T_1^k \,\&\, T_2^j \vdash E$ for every value of j). At first this result is a little disconcerting, since in an effort to objectify the problem, we have reduced it to one involving judgments of priors. What we can hope is that the priors used in this context are posteriors taken from another context and that the latter have been objectified through the weight of accumulated evidence.

The result of accumulating evidence has been investigated by Redhead (1980) under a different set of assumptions. He invites us to consider a series of refutations of the core (T_1) plus auxiliary (T_2). T_2 is replaced by T_2' to accommodate the evidence E refuting $T_1 \,\&\, T_2$; then new data F that refutes $T_1 \,\&\, T_2'$ is found; T_2' is replaced by T_2'' to accommodate F; etc. If each of the successive auxiliaries is given an initial weight of .5, and if the likelihoods of each new piece of evidence (given $\neg T_1 \,\&\, T_2'^{\cdots}$, $T_1 \,\&\, \neg T_2'^{\cdots}$, or $\neg T_1 \,\&\, \neg T_2'^{\cdots}$) are equal and substantially less than 1, then the probability of T_1 is quickly driven down toward 0 by the series of refutations. This is an interesting result, but it does not provide a resolution of the original problem.

The upshot is that we have a highly qualified success for Bayesianism: the apparatus provides for an illuminating representation of the Quine and Duhem problem, but a satisfying solution turns on a solution to the general problem of objectivity of scientific inference, a matter that will occupy us in coming chapters.

8 Conclusion

The reader does not have to share the details of the sentiments I have expressed above to be convinced that applying the Bayesian apparatus to topics like the paradox of the ravens, the variety of evidence, the role of theories in scientific inference, and the problem of Quine and Duhem leads to fruitful avenues of investigation. There are many more examples of fruitfulness that could be given. Some will be developed in chapter 4 in the context of responses to challenges to Bayesianism confirmation theory. Others can be found in such Bayesian tracts as Rosenkrantz 1981, Horwich 1982, and Howson and Urbach 1989. Franklin (1986, 1990) supplies excellent case studies of experiments in physics and makes an attempt to provide a Bayesian rationale for the strategies he sees experimental physicists using to validate their results.

4 Challenges Met

Despite or perhaps because of its successes, Bayesianism is not without its detractors. One of the most serious charges against it is that its machinery does not apply to the confirmation of universal hypotheses about an infinity of individuals, since (the charge goes) the prior and thus the posterior probability of such a generalization will be flatly 0. Three versions of this worry are examined in sections 1 to 3. Section 4 explores a different worry expressed by Karl Popper and David Miller. They argue that even when the probabilification of a hypothesis takes place, no genuine inductive support can be seen in the incremental boost in probability. Section 5 is devoted to Richard Miller's charge that Bayesianism is just as broken-backed as is HD methodology because the notorious problem of adhocing the auxiliary hypotheses that besets the latter has analogues that vitiate the former. Section 6 takes up Grünbaum's worry that Bayesianism commits its practitioners to an unbridled and implausible form of instantian inductivism. Section 7 explores the ability of Bayesianism to cope with Goodman's "new problem of induction." Finally, section 8 asks whether Bayesianism can account for the importance of novel predictions.

1 The Problem of Zero Priors: Carnap's Version

A *Carnapian confirmation function* $c(H, E)$ for a language is a conditional probability function (see appendix 1 of chapter 2) defined on pairs of sentences H, E of the language, where E is noncontradictory. From the axioms of conditional probability it follows that $c(H \,\&\, E, t) = c(E, t) \times c(H, E)$, where t is a tautology. If c is strictly coherent (see chapter 2) so that $c(E, t) \neq 0$ for a noncontradictory E, then we can write $c(H, E) = c(H \,\&\, E, t)/c(E, t)$. If we set $m(\cdot) \equiv c(\cdot, t)$, $c(H, E) = m(H \,\&\, E)/m(E)$, we see that the confirmation function is determined by the *measure function m*. (When c is not strictly coherent, the story becomes more complicated, but the details will not be rehearsed here.)

Following Carnap (1950, 1952), let us now specialize to a language L_N^K containing K monadic predicates $P_1, P_2, \ldots, P_K$, assumed to be logically independent, and N individual constants $a_1, a_2, \ldots, a_N$. A *state description* specifies for each P_i and a_j whether or not P_i applies or fails to apply to a_j. In this setting, a measure function is an assignment to the state descriptions of positive weights that sum to 1, and the $c(H, E)$ determined by this measure function is the ratio of the sum of the weights attached to the state