BAYESIANISM AND DIVERSE EVIDENCE: A REPLY TO ANDREW WAYNE*

WAYNE C. MYRVOLD†‡

Andrew Wayne (1995) discusses some recent attempts to account, within a Bayesian framework, for the "common methodological adage" that "diverse evidence better confirms a hypothesis than does the same amount of similar evidence" (112). One of the approaches considered by Wayne is that suggested by Howson and Urbach (1989/1993) and dubbed the "correlation approach" by Wayne. This approach is, indeed, incomplete, in that it neglects the role of the hypothesis under consideration in determining what diversity in a body of evidence is *relevant* diversity. In this paper, it is shown how this gap can be filled, resulting in a more satisfactory account of the evidential role of diversity of evidence. In addition, it is argued that Wayne's criticism of the correlation approach does not indicate a serious flaw in the approach.

1. Correlation. We are familiar with the notion of statistical correlation between two random variables. We can define a correlation between propositions p, q by imagining a set of data in which the points (1,1), (1,0), (0,1), and (0,0) occur with relative frequency $P(p\&q)$, $P(p\&\sim q)$, $P(\sim p\&q)$, and $P(\sim p\&\sim q)$, respectively. (In the case that some of these probabilities are irrational numbers, we imagine a sequence of increasing data-sets, with these relative frequencies as limits, and the correlation defined below considered as a limit.) This is equivalent to defining a random variable x, which is equal to 1 if p and 0 if $\sim p$, and similarly defining y to be 1 if q and 0 if $\sim q$. The method of linear regression can be applied to this data set, yielding a correlation coefficient:

$$R(p,q) = r_{xy} = (P(p\&q) - P(p)P(q))/\surd(P(p)P(\sim p)P(q)P(\sim q)). \quad (1)$$

The quantity so defined displays the expected limiting behavior: $R(p,q) = 0$ iff p and q are probabilistically independent, while $R(p,p) = 1$ and $R(p,\sim p) = -1$.

Two propositions are positively correlated if and only if the truth of one increases the probability that the other is true also; $R(p,q) > 0$ iff $P(q|p) > P(q)$, $R(p,q) < 0$ iff $P(q|p) < P(q)$, and $R(p,q) = 0$ iff $P(q|p) =$

*Received September 1995.

†The author would like to thank Georgios Asteris and an anonymous referee for their helpful comments and suggestions.

‡Send reprint requests to the author, 6 Brook Street, Sherborn, MA 01770.

Philosophy of Science, 63 (December 1996) pp. 661–665. 0031-8248/96/6304-0014$2.00

$P(q)$. Therefore, we can take as an alternate measure of probabilistic similarity the quantity:

$$S(p,q) = P(q|p)/P(q) = P(p|q)/P(p) = P(p\&q)/P(p)P(q). \qquad (2)$$

In case $P(p)$ or $P(q)$ is equal to zero, take $S(p,q) = 1$. $S(p,q)$ is greater than, less than, or equal to 1 according as p and q are positively correlated, negatively correlated, or independent. This definition is suggested by a remark of Howson and Urbach (1989, 114/1993, 160):

> The idea of similarity between items of evidence is expressed naturally in probabilistic terms by saying that e_1 and e_2 are similar provided $P(e_2|e_1)$ is higher than $P(e_2)$; and one might add that the more the first probability exceeds the second, the greater the similarity.

S was explicitly defined by Andrew Wayne (1995, 113). A similar approach to the notion of diversity of evidence has been taken by John Earman (1992, 77–79), who defines diversity of evidence differently but with the same underlying intuition: a set of evidence-statements displays a high degree of similarity if the occurrence of some of the items in the set makes the occurrence of the others highly likely.

The notion of similarity can be extended to bodies of evidence consisting of more than two propositions by defining

$$S(e_1, \ldots, e_n) = P(e_2|e_1)/P(e_2) \times P(e_3|e_1\&e_2)/P(e_3) \times \ldots \times P(e_n|e_1\& \ldots \&e_{n-1})/P(e_n) = P(e_1\& \ldots \&e_n)/(P(e_1) \times \ldots \times P(e_n)). \qquad (3)$$

We may also define a conditionalized form:

$$S(e_1, \ldots, e_n|h) = P(e_1\& \ldots \&e_n|h)/(P(e_1|h) \times \ldots \times P(e_n|h)). \qquad (4)$$

2. Confirmation. Evidence e confirms a hypothesis h if and only if $P(h|e) > P(h)$; we define the *degree of confirmation* of h by e as the ratio $P(h|e)/P(h)$. It is well-known that, if two pieces of evidence e_1 and e_2 each separately confirm a hypothesis h, their conjunction need not confirm h and may in fact disconfirm it (this was pointed out by Carnap 1950, 382–83; see also Salmon 1975, 14–15; Jeffrey 1992, 59). The question arises: under what circumstances does a conjunction of confirming evidence confirm?

It might be supposed that e_1 and e_2, each individually supporting a hypothesis h, together more strongly support h if e_1 and e_2 are independent or negatively correlated. This is not true in general, however. It turns out that the quantity which plays a key role in the degree of confirmation of h by $e_1\&e_2$ is not $S(e_1,e_2)$, but rather the ratio of $S(e_1,e_2|h)$ to $S(e_1,e_2)$. The degree of confirmation of h by $e_1\&e_2$ is a product of three factors: the degree of confirmation of h by e_1 alone, the degree of confirmation of h

by e_2 alone, and an "interaction term" which is the ratio of the degree of similarity of the body of evidence, conditionalized upon h, to its prior degree of similarity.

$$P(h|e_1\&e_2)/P(h) = S(e_1,e_2|h)/S(e_1,e_2) \times P(h|e_1)/P(h) \times P(h|e_2)/P(h) \quad (5)$$

For a body of evidence consisting of n statements $\{e_1, \ldots, e_n\}$, we have:

$$P(h|e_1\& \ldots \&e_n)/P(h) = S(e_1, \ldots ,e_n|h)/S(e_1, \ldots ,e_n) \times P(h|e_1)/P(h) \times \ldots \times P(h|e_n)/P(h). \quad (6)$$

Just as an unexpected piece of evidence confirms a hypothesis if the hypothesis renders it more probable (that is, if the probability of the evidence, conditionalized upon the hypothesis, is higher than its prior probability), a diverse body of evidence confirms a hypothesis more strongly if the hypothesis renders the evidence less diverse.

Howson and Urbach do not discuss the conditionalized form of the similarity measure, and as a result their account of the evidential role of diversity of evidence is incomplete. They write that, if e_1 and e_2 are similar pieces of evidence, "[t]his means that e_2 would provide less support if e_1 had already been cited as evidence than if it was cited by itself" (1989, 114/1993, 160). This seems to mean that if $S(e_1,e_2) > 1$, then $P(h|e_2\&e_1)/P(h|e_1) < P(h|e_2)/P(h)$. This is correct for the special case in which h entails $e_1\&e_2$. It is not true in general, however, as the following example shows.

A card is drawn at random from a deck of five cards. Let h be the hypothesis that the chosen card is #1 or #2, let e_1 be the statement that the chosen card is #1 or #3, and let e_2 be the statement that the chosen card is #1 or #4. Then e_1 and e_2 are positively correlated, since $S(e_1,e_2) = 5/4 > 1$. However, e_2 provides more support to h if e_1 has already been cited as evidence than it would if cited by itself— $P(h|e_2\&e_1)/P(h|e_1) > P(h|e_2)/P(h)$. The correct statement would be: if the prior degree of similarity of e_1 and e_2 is higher than their degree of similarity conditionalized on h, then e_2 would provide less support to h if e_1 had already been cited as evidence than if it would if cited by itself. That is: if $S(e_1,e_2) > S(e_1,e_2|h)$, then $P(h|e_2\&e_1)/P(h|e_1) < P(h|e_2)/P(h)$.

Howson and Urbach's account of the role of similarity and diversity of evidence is satisfactory for the special case in which all the evidence is entailed by the hypothesis, and this is the context in which their account is introduced. In such a case, $S(e_1, \ldots, e_n|h) = 1$, and so this factor drops out from equation (6). In the general case, the role of the hypothesis in increasing or decreasing the diversity of the evidence must be taken into account.

3. Wayne's Criticism. The degree of similarity of a body of evidence is defined in terms of prior probabilities. e_1 and e_2 are similar, in this probabilistic sense, if and only if the occurrence of one increases the probability of the occurrence of the other. Whether a particular body of evidence will be judged similar or diverse, therefore, depends on the initial assignment of probabilities.

It is characteristic of Bayesian approaches to scientific reasoning that no attempt is made to specify a unique "correct" probability assignment. Instead, one begins with some initial probability assignment, and conditionalizes from there. The initial probabilities are not assumed to be "prior" in an absolute sense; they may depend on background knowledge. The background knowledge is sometimes included explicitly: $P(\bullet) = P(\bullet|b)$ for some body of background knowledge b.

Andrew Wayne's objection to Howson and Urbach's account of the role of diverse evidence in confirmation is that it omits a discussion of how judgments of similarity depend on theoretical context.

> It was sound, before Maxwell, to regard electromagnetic and optical phenomena as diverse, while today they are judged very similar. Clearly, judgments of diversity are always made relative to a given theoretical context; what look like disparate phenomena in one context may appear closely akin in another. Without a satisfactory Bayesian account of how judgments of diversity—assignment of degree of correlation among elements of a data set—depend on the theoretical context, the correlation approach is crucially incomplete. (Wayne 1995, 115–116)

Curiously, he seems to think that the problem of assessing diversity is a distinct problem from that of assigning prior probabilities.

> Yet Bayesianism is always incomplete—it does not purport to explain our prior probability assignments, for instance. . . . My objection is not simply that the correlation approach is incomplete in this way, but that it suffers a further lack. Our understanding of variety of evidence involves intuitions about how judgments of diversity are affected by the theoretical context, intuitions which underwrite, for example, our understanding of the case of electromagnetic and optical evidence before and after Maxwell. The correlation approach remains crucially incomplete until it is supplemented by a Bayesian account of these intuitions. (Wayne 1995, 116)

On the correlation approach, similarity of evidence is defined in terms of probabilities. Once these probabilities are specified in some way, therefore, so is the degree of similarity of any given body of evidence. If the probabilities assigned take into account relevant background knowledge, which can include highly theoretical items such as knowledge of Maxwell's equations, then *ipso facto* so will the assessments of similarity. How, then, can not accounting for assessments of similarity constitute a "further lack" beyond the problem of accounting for prior probabilities?

On the correlation account, similarity is relative to a probability assignment, but, it should be remembered, so is confirmation. Equation (6) holds for any assignment of probabilities to a set of propositions. It therefore indicates a relation between the probabilistic similarity of a given body of evidence, as measured by a given choice of initial probabilities, and the degree to which that evidence confirms a hypothesis, measured by the same choice of initial probabilities, which holds independently of the choice of initial probability function. The fact that no account was given of how these initial probabilities are obtained is a virtue, rather than a defect, of this approach to the issue of diversity of evidence, as it thereby achieves generality. In any particular application, further specification of the probabilities must be made, but, however this is done, the relation between degree of confirmation and degree of similarity indicated by equation (6) will hold.

REFERENCES

Carnap, R. (1950), *Logical Foundations of Probability*. Chicago: University of Chicago Press.

Earman, J. (1992), *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*. Cambridge, MA: The MIT Press.

Howson, C. and P. Urbach (1989), *Scientific Reasoning: The Bayesian Approach*. La Salle, IL: Open Court. Second edition, 1993.

Jeffrey, R. (1992), *Probability and the Art of Judgment*. Cambridge: Cambridge University Press.

Salmon, W. C. (1975), "Confirmation and Relevance", in G. Maxwell and R. M. Anderson, Jr. (ed.), *Induction, Probability, and Confirmation. Minnesota Studies in the Philosophy of Science,* Volume VI. Minneapolis: University of Minnesota Press, pp. 3–36.

Wayne, A. (1995), "Bayesianism and Diverse Evidence." *Philosophy of Science* 62: 111–121.