ORACLES, AESTHETICS, AND BAYESIAN CONSENSUS

JEFFREY A. BARRETT†‡

University of California, Irvine

In order for Bayesian inquiry to count as objective, one might argue that it must lead to a consensus among those who use it and share evidence, but presumably this is not enough. It has been proposed that one should also require that the consensus be reached from very different initial opinions by conditioning only on basic experimental evidence, evidence free from subjective, social, or psychological influence. I will argue here, however, that this notion of objectivity in Bayesian inquiry is too narrow.

1. The Bayesian Model. On the Bayesian model of inquiry one starts with a set of hypotheses and a prior probability for each hypothesis being true. One then goes about gathering empirical evidence, and the probability associated with each hypothesis is updated on the basis of how well it accounts for the new evidence.

Let H be a hypothesis that one wants to investigate, and let K be the background beliefs that one brings to the investigation. The prior probability of H given K is written $P(H|K)$. Suppose that one takes the results of some proposed experiment to be relevant to the truth of H. $P(E|K)$ is the degree to which the evidence E is expected given one's background knowledge K, and $P(E|H\&K)$ is a measure of how well the truth of H would account for the evidence E. If one knew that E was a possible experimental result, then $P(H|K)$, $P(E|K)$, and $P(E|H\&K)$ would be probabilities that one could at least in principle determine before performing the experiment. Suppose that one in fact performs the proposed experiment and gets the result E. What posterior probability $P_{new}(H|K)$ should one assign to the hypothesis H given K after receiving the new evidence E?

Bayes's theorem and the principle of strict conditionalization (and more generally Jeffrey conditionalization) provide an explicit prescription for updating probabilities on the basis of new evidence. It follows from the definitions (or the axioms) of probability theory (depending on how one sets things up) that

$$P(H|E\&K) = \frac{P(E|H\&K)}{P(E|K)} P(H|K) \tag{1}$$

This is Bayes's theorem. The principle of strict conditionalization tells us to update our probabilities so that $P_{new}(H|K) = P_{old}(H|E\&K)$, and when the evidence is itself uncertain, Jeffrey conditionalization provides a slightly more involved prescription for how to update probabilities.[1]

The Bayesian prescription for updating probabilities has many of the properties that one would want. Bayes's theorem tells us what factor one ought to multiply one's prior probability by in order to get a posterior probability. Consider this

†I would like to thank Brian Skyrms and Ermanno Bencivenga for comments on an earlier draft of this paper.

‡Department of Philosophy, University of California, Irvine, CA 92717.

[1]See Jeffrey (1983, 169), and also Howson and Urbach (1989, 284–287).

Philosophy of Science, 63 (Proceedings) pp. S273–S280. 0031-8248/96/63supp-0032$2.00

factor: If H together with K accounts for E well, then the factor is larger than it would have been otherwise. Further, if E is surprising, if it is something that one would not have expected given K alone, then the factor is larger still. If H together with K does not account for E any better than K alone, then the experiment is irrelevant to the acceptance of H and the new evidence does nothing to change one's probability assignment. Or if K alone accounts for E better than H together with K, then the factor is less than one and the probability one assigns to H is smaller after the experiment than before. Moreover, one can argue that one would lose money if one bet on probabilities that were updated in any other way.[2] There are then compelling arguments for the Bayesian method of updating probabilities.

2. A Notion of Objectivity for Bayesian Inquiry. One can show that if all inquirers begin equally dogmatic (that is, if they assign zero probability to the same hypotheses) and if they share evidence and update probabilities according to the prescription just described, then they will conclude that their probability assignments will almost certainly agree in the long run.[3] Thus the initial assignment of priors eventually becomes irrelevant, it washes out over time. One might argue that since the objective experimental evidence eventually overwhelms whatever subjective biases might be represented in one's assignment of prior probabilities, the community of inquirers will not only reach a consensus concerning which of a set of hypotheses is best but it will reach an *objective consensus*.

But what if the community began with closely clustered priors? They might quickly reach a consensus, but the consensus would be an artifact of the clustered priors, and one might thus worry that it would not be objective. A possible response to this worry would be to require that a consensus develop from very different priors in order to count as objective. Following Smith (1986), Earman claims that "the essence of scientific objectivity" is "the emergence of an evidence-driven consensus from widely differing initial opinions" (Earman 1992, 141). If inquirers fail to reach a consensus or if they reach a consensus from closely clustered priors, then they cannot consider the results of their inquiry to be objective.

It is also important to this view that the consensus be driven by *the right type of evidence*. The evidence must be restricted to something like the straightforward reports of the results of laboratory experiments, reports that one might expect to be relatively free of subjective bias. Earman recognizes that scientists sometimes condition on other types of evidence, reports about the opinions of fellow investigators, for example, but he argues that to the extent that such evidence is responsible for the consensus reached it counts against its objectivity

> Some of the evidence on which one conditionalizes consists of reports about the opinions of fellow investigators. If evidence of this sort, rather than reports of experiments and observations played the major role in explaining consensus, then the intersubjectivity of opinion would not constitute the kind of objectivity sought but would be closer to the mob rule which Kuhn's critics read in the *Structure*. (Earman 1994, 8)

[2]See Ramsey 1926 for a Dutch-book justification of the basic axioms of probability theory; and see Teller 1973 for a Dutch-book justification of strict conditionalization, Skyrms 1987 for a Dutch-book justification of Jeffrey conditionalization, and Earman (1992, 38–44 and 46–51) for a critical discussion of the various Dutch-book arguments.

[3]See Earman (1992, 141–147) for a short discussion of the conditions for merger of opinion and some of the relevant results. The reason why everyone must be equally dogmatic should be clear from the prescription for calculating posterior probabilities—if one assigns a prior of zero, then one's posterior probability will always be zero.

In other words, an agent might condition on evidence concerning her colleague's opinions as long as the results of such conditioning, like the agent's priors, get washed out by basic empirical evidence, evidence not tainted by the fuzziness of social, aesthetic, and psychological considerations.

It is probably wise to resist mob rule, and it is certainly impressive when a community of inquirers can start with very different priors and reach a consensus by strict conditionalization on basic laboratory evidence. I will argue, however, that one might under certain circumstances condition on social, aesthetic, or psychological evidence and not worry that this might somehow taint the objectivity of inquiry, and I will argue that objectivity does not necessarily require widely differing priors.

3. Almost Anything *Might* Count as Objective Evidence. One might imagine a conference at Delphi on recent advances in particle physics where the physicists in attendance consult the oracle of Apollo during breaks between talks. Suppose that the oracle tells each scientist that a proposal made by one of their colleagues at the Friday morning session is false, suppose that the scientists take the oracle seriously and condition on this evidence, and suppose that there is consequently a consensus against the proposal by the end of the conference. Under such circumstances the proposal might be ruled out as a serious possibility by the community of particle physicists without anyone testing it empirically, at least not in the usual sense of an empirical test.

Since the consensus was reached without any direct appeal to basic experimental evidence, one might, however, suppose that it is nothing more than the expression of the subjective whim of the Priestess of Apollo (or perhaps Apollo himself) and that it must consequently fail to represent an objective consensus. But we do not know a priori what types of evidence we ought to take as relevant to the truth of a hypothesis. This too is a matter for empirical inquiry. What if Apollo's oracle had a long track record of making reliable responses when consulted on questions concerning particle physics? If this were the case, then one might suppose that the responses of the oracle would continue to be reliable, take its response to be relevant to the acceptability of the proposal, and consequently condition on it. Moreover, this might allow for a very rapid development of consensus, a consensus that anyone convinced of the reliability of the oracle, for whatever reason, would judge to be justified and presumably also judge to be objective even without understanding why the oracle was such a reliable source of evidence.[4]

Just as one might under certain circumstances condition on the pronouncements of oracles without worrying about undermining the objectivity of inquiry, one might condition on the opinions of colleagues even when the opinions are provided without supporting argument: the point is simply that a good Bayesian can and will condition on anything that she takes to be relevant to the truth of the hypothesis at hand, and what she takes to be relevant depends on her prior conditional probabilities and her experience. If she takes her colleague's judgment to be relevant to the truth of the hypothesis, then she will use it to revise her beliefs. Moreover, if the colleague has a long track record of making accurate judgments in some domain, then everyone who has this evidence and did not initially assign zero probability to the colleague being a reliable oracle will eventually condition

[4]Reichenbach (1938, 1949) argued that the inductive method would always share in the success of any successful method of prediction. If a crystal gazer consistently predicted future events, then one could inductively infer that she would continue to be successful and thus share in the gazer's success. What is true of inductive inference is also true of Bayesian belief revision.

on the colleague's judgments. The colleague's status as an oracle for the community might then help to explain the rapid development of consensus. Indeed, if the oracle's judgments were held in sufficiently high esteem, then the community might reach a very rapid consensus without any basic experimental evidence.

It is easy to find examples in the history of science that illustrate the importance of expert judgment in generating consensus. The degree to which scientists rely on such evidence, however, is perhaps best illustrated by considering the reactions of scientists when there is *disagreement* among the acknowledged experts in a field. The history of quantum mechanics provides a rich source of disagreement among experts.

When Bohr, Kramers, and Slater described their proposal for how quantum mechanics might be developed, Einstein rejected it immediately. This conflict between experts led to a general state of confusion. As Pais described the situation

> Einstein and Bohr, the two leading authorities of the day, were locked in conflict (the word *conflict* was used by Einstein himself). To take sides meant choosing between the two most revered physicists. Ideally, personal considerations of this kind ought to play no role in matters scientific, but this ideal is not always fully realized. (Pais 1982, 420)

Pauli found the conflict between authorities troubling, even though he claimed that he himself found it impossible to appeal to authority on questions of science. In a letter to Bohr, Pauli wrote

> Even if it were psychologically possible for me to form a scientific opinion on the grounds of some sort of belief in authority (which is not the case, however, as you know), this would be logically impossible (at least in this case) since here the opinions of two authorities are so very contradictory. (Pauli 1924; quoted in Pais 1982, 420)

Basic empirical evidence finally decided that in this case it was Einstein's intuitions and not Bohr's that were right.

Another example of Einstein's influence in the physics community came two years later. Within five months of the publication of Born's 1926 papers that described his new formulation of quantum mechanics, Einstein had rejected the theory. Einstein told Born that his formulation of quantum mechanics was "certainly imposing . . . but an inner voice tells me that it is not yet the real thing" (Einstein 1926, 91). Bohr liked Born's probabilistic interpretation of the wave function, but Born was nonetheless upset by Einstein's verdict. He said that

> Einstein's verdict on quantum mechanics came as a hard blow to me: he rejected it not for any definite reason, but rather referred to an 'inner voice.' . . . [This rejection] was based on a basic difference of philosophical attitude, which separated Einstein from the younger generation to which I felt that I belonged, although I was only a few years younger than Einstein. (Born 1971, 91)

Twenty eight years after he proposed his statistical interpretation of the wave function, Born won a Nobel prize for it. He said, however, that he was not surprised at the delay

> for all the great names of the initial period of the quantum theory were opposed to the statistical interpretation: Planck, de Broglie, Schroedinger and, not least, Einstein himself. It cannot have been easy for the Swedish Academy to act in opposition to voices which carried as much weight as theirs; therefore

> I had to wait until my ideas had become the common property of all physicists. This was due in no small part to the cooperation of Niels Bohr and his Copenhagen school, which today lends its name almost everywhere to the line of thinking I originated. (Born 1971, 229)

Others were also upset by Einstein's opposition to the Copenhagen formulation of quantum mechanics.[5] Pais tells of Ehrenfest's reaction to the disagreement between Einstein and Bohr, "In tears, Ehrenfest said that he had to make a choice between Bohr's and Einstein's position and that he could not but agree with Bohr" (Pais 1982, 443). The physics community generally sided with Born and Bohr. Indeed, Einstein's continued criticism of Bohr's views on quantum mechanics eventually undermined Einstein's own credibility.[6]

For many years criticism of the Copenhagen formulation was taken to imply incompetence. Bernstein tells a story about Oppenheimer that illustrates the degree of consensus eventually generated by Bohr.

> I once saw Oppenheimer reduce a young physicist nearly to tears by telling him a talk he was delivering on the quantum theory of measurement at the Institute was of no interest, since all the problems had been solved by Bohr and his associates two decades earlier. (Bernstein 1991, 63)

Recently, however, Bohr's position has lost some of its hold on the physics community. This is not the result of new basic empirical evidence, but rather it seems that it has come about because several physicists who have track records of success in answering questions in quantum mechanics have publicly asserted that Bohr was wrong. At the 1976 Nobel Conference, for example, Murray Gell-Mann said

> The fact that an adequate philosophical presentation [of quantum mechanics] has been so long delayed is no doubt caused by the fact that Niels Bohr brainwashed a whole generation of theorists into thinking that the job was done fifty years ago. (Gell-Mann 1979, 29)

The point here is simply that real scientists care about the opinions of their colleagues. While some may be embarrassed by the fact and while basic empirical evidence is almost always taken to be preferable when it is available and judged to be relevant, expert opinion nonetheless plays an important role in generating consensus in science. But if the specific expert opinion is judged to be relevant to the truth of the hypothesis being considered, then conditioning on this evidence will not be judged to represent any failure in inquiry. Moreover, if a scientist actually conditions on such evidence, then it is presumably because she judges that it is in fact relevant.

Evidence that one might take as relevant to the truth of a hypothesis is not limited to basic empirical evidence and expert opinion. Scientists often argue for positions on what appear to be aesthetic grounds. In his biography of Einstein, Pais has argued that

> Einstein was driven to the special theory of relativity mostly by aesthetic ar-

[5]Actually, it is not at all clear that there ever was a single, unambiguous Copenhagen formulation of quantum mechanics. Bohr himself seems to have adopted several mutually incompatible views of quantum mechanics over his life, and even the clearest of his statements are, I believe, too vague to allow one to figure out precisely what he had in mind. See Bohr 1949 for one of the clearest statements of his position and his disagreement with Einstein.

[6]According to Infeld, Einstein said to him more than once that "in Princeton I am considered an old fool." After quoting Infeld, Born says that by the late 1930's Einstein "was regarded an historical relic" (Born 1971, 131).

> guments, that is, arguments of simplicity. This same magnificent obsession would stay with him for the rest of his life. It was to lead him to his greatest achievement, general relativity, and to his noble failure, unified field theory. (Pais 1982, 140)

Indeed, Einstein would often appeal to the mathematical simplicity of a theory as a reliable indication of its truth, *and he believed that the reliability of such appeals was itself supported by historical evidence concerning past success and failure in the practice of science.*

> In my opinion there is *the* correct path and . . . it is in our power to find it. Our experience up to date justifies us in feeling sure that in nature is actualized the ideal of mathematical simplicity. (Quoted in Pais 1982, 466–467)

The history of the practice of science presumably *does* provide a rich source of evidence that one can condition on to determine what practices succeed and what practices fail. For his part, Einstein took this evidence to justify his belief that the mathematical simplicity of a law is a reliable indication of its truth. While Pais and others have criticized Einstein for his "excessive reliance . . . on formal simplicity" (Pais 1982, 325), especially in his later work, Pais elsewhere shows that he like perhaps most theoretical physicists also takes aesthetic evidence seriously. In discussing recent advances in field theory, for example, Pais says, "It is now conjectured that a new kind of symmetry, tantalizingly tight and elegant, supersymmetry, . . . will help to incorporate gravity" (Pais 1986, 30).

Like expert opinion, aesthetic considerations can help to generate a consensus among scientists when there is insufficient basic experimental evidence. In the case of the violation of *CP* symmetry in particle physics Franklin reports that "the vast majority of the physics community had accepted *CP* violation by the end of 1965, even though all of the tests had not yet been completed" (Franklin 1990, 152). Here we have a consensus without the basic evidence that one might expect would be necessary. Why? Because although no one wanted to give up *CP* the other theoretical alternatives *while not incompatible with the basic empirical evidence available* were considered to be "even more unpleasant" (Prentki 1966; quoted in Franklin 1990, 152–153). By 1967 the initially stated alternatives to *CP* violation had been tested, and it turned out that the aesthetic intuitions of the community were right—none of the stated alternatives to *CP* violation were consistent with the empirical evidence given what the community was willing to revise in light of the new evidence (Franklin 1990, 152–157).

One does not need to know *why* a particular type of evidence is relevant to the truth of a hypothesis in order to condition on it, but this does not mean that such knowledge is irrelevant to inquiry. Indeed, if one believed that one's intuitions might serve as reliable evidence, then one would naturally want to know why, and the plausibility of the explanation would presumably end up influencing the degree of confidence that one had in such evidence.

It is easy enough to tell a story of how conditioning on social, aesthetic, or psychological evidence might end up leading to success. One might, for example, expect that along with other creatures we have natural dispositions that routinely lead to practical success. After all, we believe that our innate dispositions were forged in an environment where those dispositions that led to failure were not as likely to be passed on to future generations as those that led to success. Something similar is also true of the acquired dispositions of scientists. The use of expert opinion, the standards of simplicity and elegance, the details of laboratory practice, the typographical formats of research journals, etc. have presumably all evolved

to facilitate the aims of inquiry since one would expect that those practices that routinely led to error were less likely to be imitated by others and more likely to be dropped by those who had mistakenly adopted them. This does not, however, mean that all is well in the actual practice of science; indeed, we have very good inductive evidence that the methodological practices of science will improve as the scientific community continues to try to identify sources of error and drop them from their practice.

Social, aesthetic, and psychological evidence might be thought of as less certain than basic empirical evidence but not fundamentally different in kind. This would explain why scientists might reach a consensus without basic empirical evidence, take the consensus to be objective, yet nonetheless prefer basic empirical evidence if they can get it.

4. Closely Clustered Priors Do Not Necessarily Count against the Objectivity of Inquiry. The problem of how worried one ought to be about closely clustered priors arises when Franklin describes a problem that he feels must be addressed by a satisfactory account of scientific inquiry.

> . . . if different scientists have different judgments as to the prior probabilities of a hypothesis, as they do, then even if they conditionalize on the evidence, they will not agree on the posterior probabilities. . . . Therefore, given these different priors a consensus will not develop. This would certainly be at odds with the actual history of science, where we do see scientists agreeing on what they view as the best theory or hypothesis. (Franklin 1990, 101)

The explanation that Franklin gives for the actual development of consensus is that "in the practice of science the estimates of the priors do not differ enough to prevent the scientific community from agreeing on the best confirmed or supported hypothesis, given a reasonable amount of evidence" (Franklin 1990, 101). But if Franklin's explanation is right, then on the Smith-Earman notion of objectivity, scientific inquiry is not objective.

It is easy, however, to tell a story where inquirers are not in any way worried about their closely clustered priors. Suppose that the scientists at Delphi believe that the oracle's pronouncements are well correlated with what will subsequently be found to be the truth and thus decide to use the oracle *to assign priors* to their colleague's proposal. Suppose that the oracle tells each scientist to assign a low prior, and that there is consequently a rapid consensus. This consensus would presumably be considered objective by anyone who believed, for whatever reason, that the oracle was reliable in assigning priors. If the oracle had a long track record of assigning accurate priors, for example, one would have little reason it seems to worry about the objectivity of the consensus from closely clustered priors.

5. Another Notion of Objectivity. If the Smith-Earman notion of objectivity is too strong, then we need another. The expectation of reaching a consensus among rational inquires is a plausible precondition for objective inquiry in a community, but it is clearly not enough.[7] While everyone may believe that given how she assigned priors she conditioned on sufficient objective evidence to participate in the consensus and while everyone participating in the consensus may believe that ev-

[7]If one requires the expectation of a consensus for communal inquiry to count as objective, then this already places a relatively strong constraint on objective inquiry, since an objective inquirer would have to follow Shimony's suggestion and assign a nonzero prior to the truth of any hypothesis or practice seriously proposed by a colleague.

eryone else did in fact arrive at the right conclusion, some may believe that others arrived at the right conclusion *for the wrong reasons*, that influences that were irrelevant to the truth of the hypothesis were ultimately responsible for the consensus, and that the community's consensus thus failed to be objective. This suggests that for a consensus to be judged objective one must believe that, given the reliability of the methods used to assign priors, the inquirers have considered sufficient objective evidence, evidence one judges to be genuinely relevant to the truth of the conclusion, to justify their participation in the consensus.

With this suggestion we have almost come full circle. If one believed that assignments of priors were arbitrary or that only basic empirical evidence was reliable, then on the notion of objective inquiry suggested here one would agree with Smith and Earman that a consensus formed from closely clustered priors or formed by conditioning on something other than basic empirical evidence would fail to be objective. But again it seems that we can only provide a Bayesian account for the rapid development of consensus in science if we take seriously the role that social, aesthetic, and psychological influences play in rational deliberation. Further, since there are circumstances where such influences would not count against reliability or intersubjectivity, such influences do not necessarily count against the objectivity of scientific inquiry.

REFERENCES

Bell, J. S. (1987), *The Speakable and Unspeakable in Quantum Mechanics*. Cambridge: Cambridge University Press.

Bernstein, J. (1991), "King of the Quantum", *The New York Review of Books*, vol. XXXVIII, no. 15: 61–63.

Bohr, N. (1949), "Discussion with Einstein on Epistemological Problem in Atomic Physics", in P. Schilpp (ed.), *Albert Einstein: Philosopher-Scientist*. New York: Harper and Row, pp. 199–242.

Born, M. (1971), *The Born-Einstein Letters*, Commentaries by M. Born, I. Born (trans.), P. Atkins (ed.). New York: Walker Publishing.

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

———. (1994), "The Rationality and Objectivity of Scientific Knowledge", unpublished manuscript.

Einstein, A. (1926), "Letter to M. Born: 4 December 1926", in *The Born-Einstein Letters*, commentaries by M. Born, P. Atkins (ed.). New York: Walker Publishing.

Franklin, A. (1990), *Experiment, Right or Wrong*. Cambridge: Cambridge University Press.

Gell-Mann, M. (1979), "What are the Building Blocks of Matter?", in *The Nature of the Physical Universe*, D. Huff and O. Prewitt (eds.). New York: Wiley.

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

Jeffrey, R. C. (1983), *The Logic of Decision*. Chicago: University of Chicago Press.

Pais, A. (1982), *Subtle is the Lord—the Science and the Life of Albert Einstein*. New York: Oxford University Press.

———. (1986), *Inward Bound: of Matter and Forces in the Physical World*. New York: Oxford University Press.

Ramsey, F. P. (1926), "Truth and Probability," in Ramsey's *Philosophical Papers*, D. H. Mellor (ed.), New York: Cambridge University Press.

Reichenbach, H. (1938), *Experience and Prediction*. Chicago: University of Chicago Press.

———. (1949), *The Theory of Probability*. Berkeley: University of California Press.

Skyrms, B. (1987), "Dynamic Coherence and Probability Kinematics", *Philosophy of Science* 54: 1–20.

Smith, A. F. M. (1986), "Why Isn't Everyone a Bayesian? Comment", *American Statistician* 40: 10.

Teller, P. (1973), "Conditionalization and Observation", *Synthese* 26: 218–258.