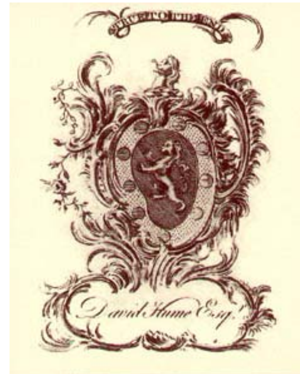

Hume's Sceptical Argument Against Reason
Fred Wilson
Hume Studies Volume IX, Number 2 (November, 1983) 90-129.

HUME'S SCEPTICAL ARGUMENT AGAINST REASON

In the section of the Treatise entitled *Of scepticism with regard to reason* Hume considers the mind as reflecting upon its own activities, monitors them as it were, and then adjusts them in accordance with certain principles and strategies.[1] What it discovers is that in drawing inferences, the mind sometimes errs. In the light of this knowledge, and in accordance with rational principles of epistemic probabilities, it adjusts its response to any inference it makes from one of certainty to one with epistemic probability of a degree somewhat less than certainty. But this rational adjustment is itself the consequence of an inference, which will be subject to the same qualifications. By repeating this process, it seems that all certainty reduces to probability and all probability to zero.

Hume's argument begins this way:

> *Our reason must be consider'd as a kind of cause, of which truth is the natural effect; but such-a-one as by the irruption of other causes, and by the inconstancy of our mental powers, may frequently be prevented. By this means all knowledge degenerates into probability; and this probability is greater or less, according to our experience of the veracity or deceitfulness of our understanding, and according to the simplicity or intricacy of the question.*[2]

It is important to be clear on the logical structure of this argument. Consider an actual case of inference. This produces a state of belief or affirmation. Hume does not say that the inference does not produce truth, i.e., a true belief. But there is the possibility it does not. Nor is this a mere logical possibility. One can form the hypothesis about causes that this process of reasoning was in fact one which introduced an element of error. One does not <u>know</u> that this hypothesis is true. But one does know that there is a <u>certain probability</u> that this <u>causal hypothesis</u> is true; as Hume puts it, it *may frequently* be true. Knowing that we have reasoned poorly sometimes (though not always) in the past, the <u>hypothesis</u> that we have on this occasion reasoned poorly has a certain probability, and we know this even though we have not verified the hypothesis for this particular occasion. On the basis of this unverified but nonetheless somewhat probable causal hypothesis we must as a matter of reason reduce the probability we attach to the belief or affirmation that was produced by the original inference. This reduction of probability is itself a matter of reasoning, however. Therefore about it, also, one can form the hypothesis that it is erroneous, and that the probability of the original affirmation was not reduced sufficiently. On the basis of this new hypothesis we must as a matter of reason reduce still further the probability of the first affirmation. And so on: as I reflect upon my use of reason, *all the rules of logic require a continual diminution, and at last a total extinction of belief and evidence.* (T183) The result here of the mind applying to its own activities what it discovers about itself is its ceasing to make any causal inferences. Feedback[3] here yields the attitude of total scepticism, in the sense of rendering it reasonable that one ought totally to suspend judgment.

To be sure, Hume does hold that no one is in fact such a total sceptic:

> *Nature, by an absolute and uncontroulable necessity has determin'd us to judge as well as to breathe and feel; nor can we any more forbear viewing certain objects in a stronger and fuller light, upon account of their customary connexion with a present impression, than we can hinder ourselves from thinking as long as we are awake, or see the*

> *surrounding bodies, when we turn cur eyes towards them in broad sunshine.* (T183)

Inferences from sample to population cannot but be fallible, and this is the best we can do, given the logical gap between sample and population. The argument for *scepticism with regard to reason* concludes not only that such inferences are fallible, that they may be wrong, but more strongly that we have the strongest of reasons for supposing that they are wrong. This is, of course, incompatible with the point Hume makes earlier in the Treatise, when he includes the section on the *Rules by which to judge of causes and effects,* that we can have good (though not infallible) reasons for supposing that certain inferences that may be wrong are in fact not wrong. Indeed, in the argument we are now looking at Hume seems to be arguing against his earlier discussion. This is, of course, not true. As Hume makes clear, his intention is quite otherwise:

> *My intention then in displaying so carefully the arguments of that fantastic sect [that is, the total sceptics], is only to make the reader sensible of the truth of my hypothesis, that all our reasonings concerning causes and effects are deriv'd from nothing but custom; and that belief is more properly an act of the sensitive, than of the cogitative part of our natures.* (T183)

which means that he must also intend to show how, upon his principles, the reasoning which leads to the *total extinction of belief and evidence* must at some point turn out to be bad reasoning, or, more accurately, to be a form of reasoning it is not rational to pursue. There are two questions. The first is: how, psychologically, is it possible that reason's *extinction of belief and evidence* cannot be maintained? And the second is: are the beliefs that destroy this *extinction of belief and evidence* reasonable beliefs? As for the first, Hume's account is perfectly straight-forward.

The process of reasoning by which belief and evidence are extinguished require a high degree of attention. A new impression then strikes us. This shifts our attention. Custom has established an association such that this impression is associated with other ideas. Given Hume's account of causation, this association is, of course, a causal belief. This association is, as a matter of psychological fact, sufficiently strong as to overcome the long train of reasonings designed to attenuate its strength. As for the second question, what it amounts to is whether the conquering association is justified by the long chain of reasonings. To this second question, the received opinion is that Hume's inference is _quite obviously_ erroneous. What I hope to establish is that this received opinion is wrong.

In Part I, I attempt to lay out clearly the logic of the probabilistic causal reasoning that Hume relies upon in his discussion. In particular, reference will be made to what Hume says elsewhere in the _Treatise_ about probabilistic causal reasoning. In Part II, I shall look briefly at two versions of the received opinion, and conclude that they go no way towards refuting Hume's argument. Finally, in Part III, I shall propose a reconstruction of Hume's argument that, one, is faithful to what Hume says; two, given certain reasonable assumptions, is probabistically sound, _i.e._, the probability _does_ tend to zero; and three, while not explicitly Hume, is historically plausible.

If I succeed in showing these things, then I shall _not_ have shown that Hume's argument that all reasoning is weak is a cogent argument. All I shall have shown is that Hume's argument is stronger than is granted by the received opinion. But _that_ will establish that Hume, on this point at least, is a better philosopher than received opinion allows, and moreover, that the argument he advances to cast doubt on reason is one that is worth refuting. Both these conclusions are worth defending.

Part I: Hume's Argument for the Destruction of Reason

In the argument against reason, Hume is explicit in taking an inference, that is, an act of inferring, to be a case of authority, something that testifies to the truth of the conclusion inferred:

> *'Tis certain a man of solid sense and long experience ought to have, and usually has, a greater assurance in his opinions, than one that is foolish and ignorant, and that our sentiments have different degrees of authority, even with ourselves, in proportion to the degrees of our reason and experience. In the man of the best sense and longest experience, this authority is never entire; since even such-a-one must be conscious of many errors in the past, and must still dread the like for the future. Here then arises a new species of probability to correct and regulate the first, and fix its just standard and proportion. As demonstration is subject to the controul of probability, so is probability liable to a new correction by a reflex act of the mind, wherein the nature of our understanding, and our reasoning from the first probability become our objects.* (T182, italics added)

Hume is not alone in making this analogy. Sextus, for one, recognizes that proof, or an argument qua actually used, is a kind of sign or testimony to the truth of what is to be proved.[4] But in any case, the analogy is there, and once it is recognized then one can apply to it the reasoning concerning chains of testimony that leads to the conclusion that the longer the chain then the lower the probability of what is testified to.

Now testimony is explicitly discussed elsewhere by Hume, most notably in his essay *Of Miracles.*[5] The crucial point about testimony is that as a rule whenever someone testifies that S then S.[6] But let us spell it out in detail.

We can say that a person a is a truth-testifier with respect to sentences s of a set S just in case that:

(1) $(s)[s \in S \supset (s \text{ is true} \equiv s \text{ is testified to by } a)]$

where, to say that s is testified to by a is to say, among other things, that if a is asked about s then a asserts s. As for the set S, this might be, for example, sentences the truth-value of which is calculated by a according to rules of arithmetic, or sentences stating what a himself has witnessed, or sentences that a has heard from an accepted authority. In any case, however, S must consist of sentences to which a has some special relation; otherwise, the "if and only if" of (1) would not work. Thus, to say that:

$s \in S$

is to say something like:

aRs

so that (1) is the law:

(2) $(s)[aRs \supset (s \text{ is true} \equiv s \text{ is testified to by } a)]$

If we restrict the range of the variable to the appropriate s in S, then (2) may be written more simply as:

(3) $(s)[s \text{ is true} \equiv s \text{ is testified to by } a]$

However, (3) does not capture the idea that it is only as a rule and not invariably that testimony is true. Sometimes even the best of truth-testifiers will mis-calculate, or mis-observe, or mis-hear, or there will be a slip in his memory, and so on. Every mind is afflicted with what Hume would call the *infirmities of reason*. (3) is thus only conditionally true, holding only when certain conditions are present, or, what amounts to the same, when the infirmities of reason are absent: in order to infer the truth of a sentence s to which a truth-testifier a testifies, we must assume that in respect to x the infirmities of reason do not afflict a. Let:

sCa

represent that in respect to s the infirmities of reason do not afflict a. This means that the law we want is not (3) but rather:

(4) (s)[s is true ≡ s is testified to by a) ≡ sCa]

or, to put it briefly:

(4^1) (F ≡ H) ≡ G

The relative frequency of S's that are G will, by (4^1), be the relative frequency that S's are such that F ≡ H obtains with respect to them. In order to find the relative frequency that F ≡ H obtains in S's, we take a sample of S's that are H, and note the percentage of these that are F. This percentage can then be used as an estimate of the probability for S's that an S that is H is also F. But for s to be H is for a to testify to s and for s to be F is for s to be true. Hence, this probability is the probability, when a testifies to something, that what he testifies to is the case. And in turn, this probability that an authority is testifying truly is the amount to which we must diminish the degree of certainty with which we affirm a sentence when that sentence is testified to by that authority.[7]

As for when the authority testifies to the truth of s and s is false, then we can "explain away" that failure by citing the presence in this case of certain infirmities of reason. If a testifies to the truth of a sentence s_1 and s_1 is false, then we have:

(5) s_1 is H

and:

(6) s_1 is F

From these and (4^1) one can infer that:

(7) s_1 is G

that states that the infirmities of reason are present in _a_ with respect to s_1. We can now _explain_ (6) _ex post facto_ using the argument:

T = (4[1])

C = (5) & (7)

E = (6)

We have positive evidence that tends to confirm (4).[8] Upon the basis of (4), we can predict the presence of _F_ (truth) on the basis of _H_ (testimony), _provided_ that _G_ is present (infirmities of reason are absent). However, often it is difficult to discover whether _G_ is in fact present or absent.[9] On some occasions we can have excellent reasons for inferring the absence of _G_, even if that fact cannot be verified by observation. This is so in the case of miracles. Here the evidence is so overwhelming that _F_ (truth) is absent that that we may reasonably infer on the basis of (4) that _G_ is absent, the infirmities of reason present. In other cases, however, we may have no independent grounds either for affirming or denying the presence of _F_. It is under just those conditions that we wish to affirm the presence of _F_ simply by predicting it using (4), and the fact of _H_ of testimony. This inference is not possible unless we know something about whether _G_ is present or absent. The basis on which to proceed is to attempt to estimate the frequency with which _H_'s are _F_'s. If the frequency of _F_'s in _H_ is such and such then, with a certainty diminished to that amount we may, upon the basis of testimony, affirm that which is testified to.

Hume discusses this general pattern of reasoning in the _Treatise_ in the section *Of the probability of causes.*[10] Hume considers cases where the same apparent cause has, at different times, contrary effects. (T132) For example, my watch works well in general and as a rule, but sometimes it does not go right. Or, an authority will in general and as a rule speak the truth, but sometimes it does not. A peasant will attribute the stopping of the watch to chance. The *artizan,* in contrast, guided by the principle that same causes have the same effects and different effects have different causes (Hume's fourth rule *by which to judge of causes and effects)* (T173)

> *easily perceives, that the same force in the spring or pendulum has always the same influence on the wheels; but fails of its usual effect, perhaps by reason of a grain of dust, which puts a stop to the whole movement.* (T132)

The artizan thus uses his causal principles to form the reasonable existential hypothesis that there is a hidden, imperceptible, cause which, when present, accounts for why the mechanism has an effect contrary to its usual effect. Often enough, the artizan's hypothesis will prove correct: he will find the speck of dust, thus confirming his existential hypothesis, and will then remove it to restore the watch to its normal state. Philosophers generalize from this sort of case:

> *...philosophers observing, that almost in every part of nature there is contain'd a vast variety of springs and principles, which are hid, by reason of their minuteness or remoteness, find that 'tis at least possible the contrariety of events may not proceed from any contingency in the cause, but from the secret operation of contrary causes.... From the observation of several parallel instances* [that is, parallel to the watchmaker's discovering the dust that is jamming the watch mechanism], *philosophers form a maxim, that the connexion betwixt all causes and effects is equally necessary, and that its seeming uncertainty in some instances proceeds from the secret operation of contrary causes.*[11]

Thus, since an authority (H) will sometimes yield contrary effects, that is, sometimes telling the truth

(F) and sometimes not, we can hypothesize a cause (G) the absence of which is a sufficient condition for the authority to be speaking the truth.

Hume also introduces probabilities into this context. He has already considered the probabilities of chances. (TI,III,XI) Take the case of a man looking at a die in a box and about to be thrown. This man concludes that the figure that appears on the most sides is the most probable. Such a one *in a manner believes, that this will lie uppermost; tho' still with hesitation and doubt, in proportion to the number of chances, which are contrary....* (T127) The crucial point in this man's reasoning is that in the causes (gravity, solidity, a cubical figure, etc.) that determine the die to fall *there is nothing to fix the particular side*, each of which is *suppos'd contingent.* (T128) Hume is thus adopting the classical definition of 'probability' based on a principle of indifference: chance arises from ignorance of causes, and chances or probabilities are equal when our knowledge is indifferent between which of the several exhaustive and mutually exclusive alternative effects will occur. There are many problems with a theory of probability based upon a principle of indifference.[12] But Hume is at least in company as good as that of Leibniz and Laplace[13] when he adopts this account of probability. Even so, Hume wants to insist that judgments of indifference or equal initial probabilities must be based upon experience. This connects his account most closely to that of the frequency theorists like von Mises than it does to such contemporary exponents of the principle of indifference as in Carnap:[14]

> *...tho' chance and causation be directly contrary, yet 'tis impossible for us to conceive this combination of chances, which is requisite to render one hazard superior to another, without supposing a mixture of causes among the chances, and a conjunction of necessity in*

> *some particulars, with a total indifference in others. Where nothing limits the chances, every notion, that the most extravagant fancy can form, is upon a footing of equality; nor can there be any circumstances to give one the advantage above another. Thus unless we allow, that there are some causes to make the dice fall, and preserve their form in their fall, and lie upon some one of their sides, we can form no calculation concerning the laws of hazard.* (T125-6)

Probability defined in terms of chance provides the model for estimating the probability that a given cause will have a certain sort of effect when, under the influence of certain hidden factors, it tends to produce contrary effects. The same basic principle that governs causal inference in general applies here:

> *'Tis evident, that when an object is attended with contrary effects, we judge of them only by our past experience, and always consider those as possible, which we have observ'd to follow from it. And as past experience regulates our judgment concerning the possibility of these effects, so it does that concerning their probability; and that effect, which has been the most common, we always esteem the most likely.* (T133)

As for the role of probability, what Hume has said in the context of chance applies equally here: *Every past experiment may be consider'd as a kind of chance; it being uncertain to us, whether the object will exist conformable to one experiment or another: And for this reason every thing that has been said on the one subject is applicable to both.* (T135) But it is always in conformity with the general causal principle of "same cause, same effect":

> *Thus upon the whole, contrary experiments produce an imperfect belief, either by weakening the habit, or by dividing and afterwards joining in different parts, that* <u>*perfect*</u> *habit, which makes us conclude in* <u>*general*</u>*, that instances, of which we have no experience, must necessarily resemble those which we have.* (T135)

We may summarize Hume's account of *scepticism with regard to reason.* On this account what the total sceptic does is apply certain lawful facts about testimony to his own reason taken as testifying to the truth of the conclusions it affirms. What then begins to happen is just what happens with respect to hearsay in contrast to direct testimony: the greater the number of steps it is from what is to be affirmed the less is the degree of certainty with which that affirmation can be made. The degree of certainty tends to diminish towards zero: all "evidence and belief" is extinguished. But not even this can be affirmed with certainty. For one has arrived at this conclusion by cases of reasoning, and reason may apply the *same* sort of reasoning to these cases. In this way, the beliefs which constitute the very foundation of the reasoning which "extinguishes" all "evidence and belief" are themselves extinguished. All reasoning is undermined. All reasoning is weak including the reasoning that all reasoning is weak.

Part II: Two Critics of Hume

In order to defend Hume's argument as strong, if not sound, it will be necessary to establish the cogency of the probabilistic reasoning that he uses. Before turning to this task, however, it will pay, I think, to show that we can already see that defenders of the received opinion, that Hume's argument is worthless, in fact quite miss the mark. I begin with H.A. Prichard.[15]

Prichard treats Hume as a sceptic. But he finds that the Treatise, so understood, is tedious, and that it makes him not sceptical but angry.[16] To be sure, Prichard thinks that much of it is clever, but it is the cleverness of ingenuity and perversity. For myself, I should say that this response should suggest to Prichard that his reading of the Treatise is incorrect. Perhaps the arguments appear perverse because they are construed as aiming at a sceptical conclusion; if construed as directed at a non-sceptical position, then perhaps they might not look so perverse or merely ingenious! However, Prichard is so convinced of the rationalistic position with respect to knowledge;[17] that he cannot but read Hume as a sceptic -- which, of course, Hume is, at least with respect to the objective necessary connections of the rationalist. But that does not mean Hume is a pyrrhonist with respect to our knowledge of causes:[18] to be sceptical of objective necessary connections is not to be a sceptic with respect to causation[19] -- though it is to be a fallibilist about the latter; but, again, to be a fallabilist is not to be a sceptic, or, at least, a pyrrhonian sceptic.

For the sceptical argument that "extinguishes" all "belief and evidence" Prichard raises three objections.[20] He asserts, in the first place, that it confuses the limit of a series with a term of the series. This is no doubt so. The limit of the series of probabilities is indeed zero, but we never actually reach that limit. All we ever reach is a fraction; indeed, if we go on long enough we can reach any fraction as small as we choose. Still, all we ever reach is the vanishingly small and not zero. We must recognize, however, that this mistake hardly modifies Hume's point: "Vanishingly Small" will do as well as "zero" in undermining belief and evidence. Moreover, given that no mathematician of that age was ever clear on the distinction between zero and the vanishingly small, that is, on the nature of infinitesimals, it is a harsh judge indeed that will condemn Hume for a similar failure of understanding. In fact, we see in Prichard's objection nothing more than a simple failure to approach Hume with the sympathy with which one should approach any philosophic classic.

Prichard objects, second, that an infinite regress cannot arise:

> *...if we judge that the faculty by which we, considering the nature of two and two, judge it to make four is infallible only to the extent of three-quarters, we inevitably are judging as a* matter of certainty *that two and two is* probably four to the *extent of three-quarters, and we cannot then take* further *account of the fallibility of our* faculties, *for we have already taken full account of it.*[21]

But clearly, one has taken full account of the fallibility of our faculties only if the probability judgment is a matter of certainty. Prichard tells us that "inevitably" the latter is so. Alas, he gives us no reason for thinking it to be so, and less for thinking that it is inevitable. All we have here is Prichard the rationalist asserting that Hume the fallibilist is wrong; but to assert is not to argue.

Prichard's third point is similar.[22] The initial judgment must be taken as being certain, since by hypothesis it is rendered probable only by reflecting on the fallibility of our judging faculty. However, if it is certain, then there is no room for doubt; if it is certain, then it cannot be arrived at by fallible faculties. Hence, Hume cannot use reflections on the latter to cast doubt on the former. However, this objection goes through only if certainty is a guarantee of infallibility. In the psychological sense of 'certainty', i.e., felt certainty, or as Hume would analyze it, maximum vivacity, there is no guarantee of infallibility. There is such a guarantee only if there is some objective necessary connection between the object of judgment and the feeling of certainty.[23] But Hume has argued in detail against objective necessary connections; the connections between objects or impressions and ideas are causal and these causal relationships are all to be given a Humean analysis.[24] Prichard's objection is successful if and only if there is such a necessary connection between objects and certainty will guarantee that judgments of the latter sort are infallible. But he does not argue for such a connection; he simply assumes it. Again all Prichard has succeeded in doing is to express his rationalistic dogmatism, while at the same time ignoring Hume's argued case that all our faculties are fallible.

Prichard may be correct in holding that the arguments of the Treatise are ingenious and perverse. But at least the Treatise gives arguments. Prichard does not even do Hume that courtesy.

Another recent critic of Hume's sceptic's argument is MacNabb, who tells us that:

> The argument is plainly sophistical but phrased in such vague terms that it would be a lengthy task to set out all the various possible cases covered and explain the fallacy in each. Take one case. I judge (1) that Bucephalus will probably win the 2:30. But then I reflect (2) that I am not a good judge of form. Then I reflect (3) that I am not a very good critic of my own performances in these matters. The force of (3) is to counteract (2) and leave (1) unchanged but subject -- as it always should have been -- to the proviso "unless I am mistaken." And how am I supposed to proceed further in this regress? Where shall I find evidence of my powers of criticizing my own criticisms of my judgments, distinct from the evidence of my powers of criticizing my own judgments? I have reached an assessment of my powers of picking up my own mistakes, and beyond that it is not possible to go. Every further step is the same step repeated.[25]

However, it is (2) rather than (3) which requires us to put (1) under the proviso "unless I am mistaken".[26] (3) does not counteract (2), but rather in turn subjects the latter to a similar proviso. (3) can then be made subject to the same proviso. And so on: the regress is clearly there. McNabb tries to break it off by arguing that the regress requires one to distinguish (a) powers of criticizing one's criticisms of judgments from (b) powers of criticizing judgments, where in fact (a) and (b) are not distinct, since evidence for (b) counts as evidence for (a). This is so: one can succeed in distinguishing (a) from (b) only by failing to note that the "criticisms of judgments" mentioned in (a) are themselves judgments. Thus, the evidence for (a) does count as evidence for (b) and, more specifically, evidence for either is evidence for "All my judgments [including critical judgments!] are fallible" or to "Each of my judgments is probably to degree $\underline{m}/\underline{n}$ mistaken". And once one has this rule then it can be applied at each stage. It is, to be sure, "the same step repeated", but nonetheless at each step the probability decreases.

In fact, Hume's argument is <u>not</u> phrased in vague terms. He himself, as we saw, places it in the context of an analogy with a chain of testimony. Once this is seen then it is not hard, I think, to generate Hume's sceptic's regress of probabilities, through some elementary applications of the probability calculus. Critics like Prichard and MacNabb are simply too hasty to judge Hume a sophist to read him carefully, let alone try to construct a reasonable interpretation in which a regress of probabilities can validly be generated.

<u>Part III: Hume's Argument Reconstructed</u>

Hume's argument against reason relies upon probabilistic causal reasoning. To defend it one must attempt to show that such reasoning is cogent. To do this will require us to invoke the probability calculus in a way that Hume never did. The reconstruction will have to be faithful to what Hume does say, that is, the argument as we have sketched it in Part I, but since the reconstruction will add details that Hume never provides, it follows that we cannot show that Hume reasoned as our reconstruction proposes. The best that can be done is to provide a little historical context sufficient to establish that the reconstruction is not only faithful to Hume but is also not wildly anachronistic.

Let us begin, then, by recalling that Hume's argument construes an inference, i.e., an act of inferring, as testifying to the truth of its conclusion, and by noting that a similar argument about chains of testimony and diminution of probabilities occurs in Bentham's Rationale of Judicial Evidence where he employs what is essentially Hume's argument in order to exclude hearsay evidence:

> All modifications of unoriginal evidence that are of the nature of, or bear similitude to, hearsay evidence ... have this in common, -- that for every remove (mendacity and fraud out of the question) they afford an additional chance of incorrectness and incompleteness.[27]

and he continues a little later:

> Supposed extra judicially stating or narrating witnesses may have stood in a series of any length, one behind another. The causes of untrustworthiness applying to every human being, and, to every being of which nothing more is known than that he or she is human, with equal force, -- it is evident that, the longer the line of these supposed witnesses, the less is the probative force of their supposed testimony.[28]

Now, Bentham is acknowledged to be the most significant theorist of legal evidence. As one authority puts it, "Among the theorists Bentham stands first, though possibly not the first in time...."[29] That by itself establishes, I think, that Hume's argument is not as implausible as such critics as Prichard and McNabb suggest. On the other hand, Bentham, like Hume, provides no reasoning explicitly in terms of the probability calculus. The task remains, then, of providing the reconstruction we are aiming at. But Bentham does provide a reminder that testimony and probabilistic considerations have long been intertwined, and a suggestion that it is in this context that we should attempt to reconstruct Hume's -- and Bentham's -- reasoning.

To this end we may start by noting what happens to the ancient distinction between _scientia_ and _opinio_ under the impact of the criticisms of Locke and Hume.[30] _Scientia_ consisted of infallible and certain knowledge of objective necessary connections. _Opinio_ consisted of fallible belief based primarily on authority.[31] Locke and Hume systematically criticized the traditional view that there are objective necessary connections, that knowledge consists of grasping these, and that demonstrative syllogisms beginning from infallibly known premisses lay out the ontological structure of the world. After this critique neither _scientia_ nor the demonstrative syllogism of _scientia_ remains. But neither does _opinio_ remain. Or, at least, the dialectical syllogism, which (together with rhetoric) dominated the theory of _opinio_, comes to be but a minor part of case-making, that is, case-making with respect to what the facts are and how we ought to understand them. To be sure, it is an essential part. The rules of formal logic remain to ensure consistency. Nor can they be eliminated, since nothing else is a test of consistency. But they come to be but a part of the methodology of empirical science, the logic of truth. The certainty of intuition and demon-stration in _scientia_ have gone. The only certainty that remains is such certainty as empirical research can yield. Gone, too, is the probability of _opinio_. Authority, what justifies premisses as probable, is no longer taken as given. The worthiness of an authority becomes just another empirical hypothesis, to be tested like all others by empirical research. Hacking has quite correctly indicated the origin of the probability within the context of opinio, attached there to the notion of acceptable propositions.[32] What he has failed to spell out is that in the context of opinio probability is connected with a magical idea of authority, be it that of the scholastics, the Bible, the Fathers, and Aristotle, or that of the humanists, the purified texts of the ancients, or that of the common lawyers, the sworn testimony of witnesses.[33] What happens as this magical context erodes is that probability -- and certainty also -- comes to be connected with the idea of the worth of empirical evidence with respect to hypotheses about matter-of-observable-fact regularities. Knowledge is no longer scientia but comes to be justified certainty about empirical regularities. And where such certainty is not possible we rely upon what Locke called judgment, and, where the matter is put to us verbally, assent or dissent.[34] Judgment or assent is probable[35] just in case it conforms with our past experience and "constant observation"[36] or conforms to the testimony of others where, however, the latter is to be evaluated as any other empirical hypothesis.[37] The syllogism is the test of consistency: "This way of reasoning [the syllogism] discovers no new proofs, but is the art of marshalling and ranging the old ones we have already."[38] Probability is part of the logic of truth: "...as the conformity of our knowledge, as the certainty of observations, as the frequency and constancy of experience, and the number and credibility of testimonies do more or less agree or disagree with it, so is any proposition in itself more or less probable."[39] Locke is content to let it go with "constant observation," no doubt trusting the basic rationality of men to normally judge well, if not always correctly. Hume gave more details. Constant observation often tends to confirm contrary hypotheses. Since confirmed, all are not unworthy of assent, but since contrary not all can be true. They are therefore to be characterized as probable.[40] The task of research is to move from such probabilities to certainty. The rules are the *Rules by which to judge of causes and effects,* (TI,III,XV), which are, in essentials, the rules of eliminative induction. Bacon's new organon is the logic of truth by which we move from probability to empirical certainty (i.e., insofar as it can be had).

But why should the "contrary causes" or contrary hypotheses be called "probable"? Here we must look more carefully at the concept of probability.

A mass event consists of a characteristic being exemplified in each member of a large group or mass of individuals or points. All persons in a country at a given time constitute a mass event; so does the set of tosses of a coin. A second characteristic may be exemplified by the individuals in this mass event. Thus, dying before a certain interval has elapsed is such a characteristic for our first mass event, coming up heads is one for our second mass event. The relative frequency of the exemplification of the second characteristic in the mass event exists. That a certain characteristic is exemplified with such and such a relative frequency in a mass event is itself a group characteristic of the mass event. Some sorts of mass events are serially ordered, but not others. Populations at a time are not, coin tossings are. Where there is a serial order, and if the exemplifications of the characteristic defining the mass event and the exemplifications of the second characteristic satisfy together certain conditions (axioms), then one has a random sequence. Where there is a serial order, one has a sequence of mass events (the first n tosses, the first n + 1 toss-es, and so on), and therefore a sequence of relative frequencies. Where there is a random sequence then the probability of the second characteristic relative to that which defines the mass events is the limit of the series of relative frequencies as the size of the mass events increases to infinity. Consider a certain individual coin being tossed. The sequence of events, this coin being tossed at t_1, this coin being tossed at t_2,..., is serially ordered. This sequence is random. The characteristic of being tossed at some time or other defines a sequence of mass events. Within these mass events the characteristic of coming up heads appears with a certain relative frequency. We form the hypothesis about the sequence that the limit of the relative frequency as the sequence tends to infinity is 1/2; i.e., that the probability of heads is 1/2. This hypothesis, if true, is a lawful regularity about the tosses of this particular coin. This hypothesis involves mixed quantification. To say p is the limit of the series of frequencies is to say that for all ϵ, there is a number N, such that for every $n > N$, where n is the (ordinal) number of the term in the series, the difference (neglecting signs) between f_n (the frequency which is the nth term) and p is less than ϵ. Such a lawful regularity enables us to assert that if the sequence of tosses were to be extended indefinitely then the relative frequency would (in the indicated sense) converge towards 1/2 as a limit. The hypothesis is an indicative prediction asserting a matter-of-fact regularity about the tosses of a single coin. The hypothesis can be generalized further, to cover the tosses of all well-balanced coins. The notion of probability for these cases is a well-defined notion, one which in no way violates the empiricist criterion of meaning. There is nothing -- or ought to be nothing -- in such a criterion which renders unacceptable the idea of extrapolating to limits.[41] The only problem is that of confirmation. Such extrapolations will, in general, given the notion of limit, involve mixed quantification. The relevant laws will, therefore, be neither conclusively confirmable nor conclusively falsifiable. In the case of probabilities the problem is to figure out what probability a sequence or each of a set of sequences has, since any initial segment we in fact observe is compatible, given the logic of limits, with any probability. But statisticians have in fact developed reliable evaluation procedures.

What must be emphasized is that such probability statements are statements of matter-of-observable-fact regularities. To be sure, they are of a logically complicated sort, but that does not render them any the less statements of regularities. This logical complication makes the use of induction in their case more precarious than it is in the case of non-probabilistic laws. "But," as Bergmann once put it, "to be precarious in this factual sense does not mean to be precarious or problematic in a philosophical sense."[42] Furthermore, these regularities may or may not be deducible from non-probabilistic laws. Whether or not deterministic explanations of probabilistic laws are possible is a question of fact. Laplace,[43] and before him Hume,[44] thought such laws as those of coin tossing could be explained deterministically; we now know this not to be so, that they are irreducibly statistical, though they can, in a way, be fit into the basic theoretical framework of classical mechanics.[45]

The coin-tossing and dice-rolling cases are the paradigm cases of random sequences of events. There were in fact problems with defining mathematically the notion of random sequence. Von Mises' idea was unsatisfactory, and Wald and others worked to refine it, without, however, providing a perfectly satisfactory notion.[46] Kolmogorov has developed a different line of research into the notion of random sequence.[47] As he has emphasized, it is crucial to the idea of applying any mathematical theory of probability.[48]

In the coin-tossing case, two outcomes are equally probable. This is the special case of the more general idea of a lottery, in which tickets are drawn from a drum (or balls from an urn). In the latter, the probability of drawing each ticket is assumed to be equal. If there are N different tickets, then for each ticket the probability of its being drawn during a long series of draws is 1/N. Just as we have found certain sorts of coin are as a matter of fact such that in a long series of tosses heads and tails are equally probable, so we have found certain sorts of ticket drawing situations are as a matter of fact lotteries in which each outcome is equally probable. Thus, we can assume certain sorts of ticket drawing situations are as a matter of fact lotteries in which each outcome is equally probable. But we can also assume certain outcomes are more or less probable than others. Thus, we can assume certain sorts of tickets are more likely to be drawn than others, or that the probability of heads is some value other than 1/2. Huygens was able to reduce the treatment of these cases to the case of the fair lottery.[49]

Taking coin-tossing and lotteries as paradigms, it becomes possible to extend the notion of probability to the case where, because events are not serially ordered, we have only frequencies, not probabilities.[50] The fact that a certain percentage of a population of size N, taken at time t_1, is male (M), the remainder female (F), has, by itself, nothing to do with probabilities. One simply has a mass event -- the population at $\underline{t}_1$ -- with another characteristic, M, occurring with a certain frequency in it. The connection with probability goes as follows. What is crucial is that the mass event is an event in a second-order mass event. In the present case, this second-order mass event consists of a group of populations of size N. In each of the mass events in the second-order mass event, M occurs with a certain relative frequency. The possibilities are zero M, one M, ..., NM; i.e., there are N + 1 possible relative frequencies. Now write each of these relative frequencies on a ticket to be put into a lottery. We can now draw tickets and obtain a series of relative frequencies. Assume the first-order mass events can be serially ordered. In the present case, we consider successive populations of size N. Then, that order corresponds to the series of relative frequencies resulting from the lottery draw provided that we make the appropriate factual assumptions of relative frequencies (male to total population) in the second-order mass event. If we assume the frequencies are uniformly distributed in the second-order mass event, then we are assuming each ticket in our "frequency lottery" is equally probable. If we assume other distributions then we are assuming certain unequal probabilities for different tickets in our "frequency lottery." We can reduce this case to the coin tossing case as follows.

We generate the same sequence of relative frequencies in another way. One considers a coin with M on one side and F on the other. This is repeatedly tossed. This sequence is divided into segments of length N. These are, of course, serially ordered. For each segment, M occurs with a certain relative frequency. These frequencies in segments of length N correspond to the relative frequencies in the non-serial mass events and to the relative frequencies yielded by the lottery draws. If we assume the tosses are independent, and that the outcome M has probability p, then the probability of nM and (N - n) F is given by:[51]

(a) $C^N_n p^n (1 - p)^{N-n}$

Thus, to make an assumption about p in our sequence of coin tosses is equivalent to making the mentioned factual assumption about the distribution of frequencies in the second-order mass event. In this way we can speak of probabilities where, strictly, we should speak of frequencies.

Arbuthnot used this analogy to coin tossing when he discussed population statistics and their stability.[52] He assumes the M-F coin is fair, so p = 1/2. He correctly calculates, using the binomial coefficients, the probabilities of frequencies in segments of length N. (a) represents the nth term of the binomial expansion of $(p + q)^N$ where q = 1 - p. (C^N_n is the binomial coefficent of that term.) If we take n to be a "random variable" and (a) to be a function of it, then we have a binomial distribution of n. Thus, Arbuthnot makes the matter of fact assumption that the distribution in the second-order mass event of frequency of males in the first-order mass events (populations of size N) is distributed binomially. This is equivalent to assuming the probabilities in our "ticket lottery" are distributed binomially. More specifically, Arbuthnot assumes a binomial distribution in which p = 1/2. If p = q = 1/2, then (a) is equal to:

$C^N_n \times 1/2^N$

As N increases in size the probability of getting any particular number n, i.e., any particular relative frequency n/N of males, becomes smaller and smaller.

Arbuthnot correctly calculates that this probability becomes smaller as N increases, and in particular that the probability of an equal number of males and females becomes very small. Equally correctly he infers that the probability of obtaining an extreme outcome never vanishes. Arbuthnot -- incorrectly -- infers that the roughly equal numbers of M and F is highly improbable, and its occurrence therefore not a "matter of chance" but due to Providence. What Arbuthnot does not ask in quantitative terms is what the probability of obtaining approximately the same number of males as females. Bernoulli's limit theorem[53] establishes that, as N becomes large, the probability tends to unity that the frequency n/N diverges only slightly from p. Thus, given Arbuthnot's distribution assumption, that $p \doteq 1/2$, the probability of obtaining populations approximately 1/2 males tends to unity. And so, contrary to what Arbuthnot invalidly inferred, the more or less constant regularity of males and females can reasonably be expected even as a "matter of chance". Actually the proportion of males to population tends over time to be slightly greater than 1/2. Nicholas Bernoulli calculated[54] that the observed relative frequencies are what one could expect if $p = 18/25$, that is, if one assumed a binomial distribution slightly different from that assumed by Arbuthnot.

Bernoulli's limit theorem leads to the prediction that if a population is large and relatively constant in size, then the proportion of males will remain relatively constant. This is, of course, an empirical prediction. It presupposes two empirical assumptions. One has already been mentioned. This is the distribution assumption. The other is that, just as the outcomes of the coin tosses are as a matter of fact independent of each other, so this person being M must be independent of that person being f, or, what is the same, the sex of the next to be born must as a <u>matter of fact</u> be independent of the sex of the just born. These factual assumptions are generalities. Their acceptance is therefore justified in the same way all other empirical generalities are justified, by their predictive success. Only, they are generalities involving mixed quantification -- otherwise they could not connect up with probability notions! -- and so what we deduce as the prediction is an existentially quantified statement which cannot be falsified. The problems of putting them to the test are essentially those of putting the probability laws of coin tossing to the test. They are problems which are technical ones in statistics rather than philosophical.

The use of statistics can be extended from these cases to the investigation of causal laws. Suppose we know that being *F* sometimes causes *H*, sometimes not. This is a very simple instance of "contrary causes". In such a case it might be reasonable to suppose some unknown factors exist such that when these are present *F* causes *H*, when they are absent *F* does not cause *H*. Call these unknown factors $\bar{G}$. (The bar indicates that we have a definite description of the type of factor, not the name of the type itself.) We are assuming, then, that the presence of $\bar{G}$ is necessary and sufficient for *F* to cause *H*, or, in symbols:

(1) $(x)[\bar{G}x \equiv (Fx \equiv Hx)]$

or, equivalently

$(x)[Fx \equiv (\bar{G}x \equiv Hx)]$

(If this looks like (4^{1}) from part I, then it is intended to!)

If *F* occurs in *a* followed by *H*, we have

(2) *Fa*

and

(3) *Ha*

from which we can deduce

(4) $\bar{G}a$

Given (1), we can deduce $\bar{G}$ is present when H is consequent upon F, even though we cannot independently identify $\bar{G}$. We can now use this information to explain the presence of H, in the pattern with which we are by now familiar from the preceding section.

The explanation is this:

(E)
(Law: (1)
(
(Initial Conditions: Fa & $\bar{G}a$
(
(Hence: Ha

Thus, where a medicine brings about recovery some but not all the time we generally infer its not working is due to the absence of certain physiological factors normally present. Consider the total population and its mutually exclusive subpopulations of size N. $\bar{G}$ occurs with some relative frequency in each of the subpopulations. If we assume a binomial distribution of these frequencies then we can, as discussed just above, equivalently speak of the probability of an individual being $\bar{G}$. Since $\bar{G}$ is present in an individual if and only if F ≡ H is also present, it follows that if f is the relative frequency of $\bar{G}$ in a subpopulation, then f is the relative frequency with which F causes H in that subpopulation; and it follows further with reference to the total population that the probability that an individual is $\bar{G}$ is the same as the probability that an individual F causes H. If we pick out a subpopulation and bring it about that each member is F, then the frequency of H in that subpopulation will be f if and only if the relative frequency of $\bar{G}$ in that subpopulation is f. The relative frequency of H's in a subpopulation of f's yields the relative frequency of $\bar{G}$'s in that subpopulation. If we pick several such subpopulations of F -- randomly, from the total population, so that independence is assured -- then we can use these data to infer statistically the probability of _H_'s in the subpopulations of _F_'s or, what is the same, the probability of $\bar{G}$'s in the total population. On the basis of such probability estimates we can use the technique of Huygens for adjusting our beliefs and actions, that a gamble is worth the mathematical expectation of that gamble,[55] and adjust our degree of assent that this _F_ will be _H_ to the probability in the total population that being _F_ causes _H_.

Obviously, we would prefer to know what characteristic it was that the definite description $\bar{G}$ denotes. Then we could identify those cases where _F_ will bring about _H_ and those cases where it won't; we would not then have to rely upon statistics. Statistics here are the mark of imperfect knowledge, our not knowing the causes we have reason to believe are operating. Probability thus comes to be used -- and, given certain factual assumptions are fulfilled (independence, etc.) -- successfully used -- where _ex post facto causal_ explanations such as (E) are available and where we also have reasonable hope to discover the means (by identifying the property $\bar{G}$) necessary for deterministic predictions.

The case described by (1), where $\bar{G}$ is unknown, is the simplest case of the operation of "contrary causes", which Hume discussed.[56] Sometimes _F_ brings about _H_, sometimes not. The idea is obviously generalizable. We could, for example, have _E_ producing H_1 sometimes and sometimes H_2. We could then hypothesize after the fashion of (1) that F & G_1 causes H_1 and that F & $\sim G_1$, or perhaps F & G_2 causes H_2. Statistics would proceed in much the same way.[57] _It is in this fashion that we come to speak of the probability of "contrary causes" and of "contrary_ effects". Such a use of 'probability' marks an imperfection in our knowledge, a gap in our knowledge of causally relevant characteristics. It is then the task of the experimental method, of the methods of eliminative induction, to fill those gaps in our knowledge, find out specifically what the characteristic $\vec{G}$ is, and so eliminate the need for statistics. However, action often requires us to use what knowledge we have, imperfect as it is, before research can be done to remove its imperfections. In that case we must rely upon the imperfect statistical knowledge -- where, one must note, to be imperfect or "gappy"[58] is not therefore to fail to be knowledge. When we do act upon statistical knowledge we proceed upon the basis of several matter-of-fact assumptions. On the basis of these judgments we can proceed to calculate the probability of kinds of particular events. Our expectation of a particular event of such a kind is proportioned to the probability, and that is to say that it is to that probability that we proportion the degree to which we assent, prior to observation, to the proposition that a particular event will be of that kind. In this context, epistemic probabilities (rational degrees of assent) are proportioned to and presuppose aleatory probabilities. Such rational adjustment of belief to probability presupposes antecedent acceptance of certain inductive generalizations.

Now consider these factual assumptions upon which the expectations and degrees of assent are based: these assumptions are inductive generalizations. Of these, there are essentially two. The one is that $\bar{G}$ exists. The other is that it has a certain distribution. The latter is a matter of mixed quantification, and partakes of the general precariousness of statistical generalizations. In the present case it further depends upon the other assumption that $\bar{G}$ exists. It is the latter assumption which is crucial. What reason might we have for believing $\bar{G}$ exists, *i.e.*, for believing that definite description is successful? That amounts to asking what reason we might have for believing a law of the sort (1) obtains even when we do not know specifically what $\bar{G}$ is and therefore could have no confirming instances upon which to base our judgment. The answer can only be: inductive. We might have some more general theory, confirmed in other cases, that enables us to deduce in the present case that (1) is true. Thus, if *F* is a drug to cure (*H*) a very specific type of disease (*e.g.*, swine flu), then a more general theory about a genus of disease (influenza), of which that of our concern was a species, might very well lead us correctly to affirm (1), without specifying specifically, only generically, what are the conditions $\bar{G}$. Or, one might rely upon some very general principle of determinism, as Hume did (T132) when considering this very sort of case.

A special case of using statistics was of special importance to early investigators. This was the case of the worth of testimony. That this should be important in this context was thoroughly natural, since, after all, probability is, as Hacking argues, a concept which does derive from the context of *opinio*, the dialectical syllogism, and the reliance upon testimony.

Assume

Fx = *x* asserts at *t* some proposition or other

and

Hx = What *x* asserts at *t* is true

i.e., The proposition *x* asserts at *t* is true

and suppose, for the moment, somewhat simplistically, that a law of the sort (1) holds for this *F* and *H*. The absence of $\bar{G}$ at some times rather than others would then explain why a person sometimes asserts what is not true.[59] Now, as above, we can use (1) or, actually, (1) quantified over both x and t, to estimate the probability in a sequence of assertings that what is asserted is true.[60] Our degree of assent to the statement that this F is H is to be proportioned to this probability. Given that a asserts something at t_1 is true. But if:

the proposition a asserts at $t_1 = S_1$

then we are, upon a's say-so, assenting to the indicated degree to the proposition that p. The probability of H in F is the probability of $\bar{G}$ in the total population of all individuals at all times. It is this probability to which we proportion our assent to S_1, and which measures the reliability of a's testimony, the extent to which we may reasonably bet upon its correctness.

Suppose now, that:

$S_1 = Fb$

$= b$ asserts S_2 at t_2

Given p, we may, as before, assent to q in proportion that $\bar{G}$ is probable in the total population. But we are not given S_1 categorically; we assent to it only to a certain degree. To what degree then should we assent to S_2? Obviously we may reasonably expect S_2 to be true just to the extent that it is probable that we draw from the population two $\bar{G}$'s in a row. Now, according to (a), this probability:

$$C^2_2 p^2(1 - p)^{2-2} = p^2$$

Since $0 \leq p \leq 1$, we will always have $p^2 < p$ unless everyone exemplifies $\bar{G}$ at all times or no one does at any time. Since, as a matter of fact, none of us tells only truths and none only falsehoods, it follows that the degree of assent we should attach to S_2 is, in the circumstances, always less than that we should attach to $\underline{S}_1$. This is what lies behind the courtroom evidentiary practice of prohibiting hear-say evidence.[61] On the other hand, consider the probability of obtaining at least one $\underline{\bar{G}}$ in two draws. By (a), that is:

$$C^2_1 \, p \, (1 - p)^{2-1} = 2p \, (1 - p)$$

The probability 2p (1 - p) is greater than p just in case $p \geq 1/2$. So, if it is more likely than not that a person is $\underline{\bar{G}}$, then the probability of getting at least one $\underline{\bar{G}}$ in a draw of two is greater than the probability of $\underline{\bar{G}}$. If, therefore, we take two witnesses the probability that the testimony of at least one is correct is greater than the probability when making a single draw that the witness' testimony is correct.. We should therefore attempt to secure several independent witnesses to an event: the probability at least one is correct is greater than the probability any is correct. So testimony derived from several independent sources is better than testimony from a single source -- provided that, of course, the probability a person is $\underline{\bar{G}}$, i.e., a truth-sayer, is greater than 1/2.

Jeremy Bentham was effectively the first[62] to provide any extended treatment of the rules of evidence in legal situations.[63] He made use of such inferences as the above, though using only relatively crude notions of probability.[64] George Bentham used relative frequencies to provide some analyses, but not probabilities.[65] J.S. Mill recognized the relevance of probabilities.[66] He also recognized the relevance of choosing a population and estimating p for that population. To attack the credibility of a witness is to try to place him in a population with respect to which the probability of $\underline{\bar{G}}$ is low, i.e., with respect to which one assents to what is testified to only with a relatively low degree of assent. Hume's discussion of miracles (EHU,X) is, as Hacking has pointed out,[67] an exemplification of Hume's ideas on probability, applied specifically to the case of testimony. His aim is, of course, to attack the credibility of any witness who testifies to a miracle. It was these same general principles that Jeremy Bentham used to rationalize the rules of legal evidence. In an important essay Waldman has argued that "...the need for a systematic treatise on evidence was most pressing after a philosophical theory of evidence regarding the practical affairs of men had been stated." [68] The philosophical theory Waldman has in mind is that of the "mitigated scepticism" of Glanvill, Wilkins, and so on, in which the ideal of absolute certainty comes to be replaced by that of moral certainty, and in which a proposition not absolutely certain remains acceptable so long as it is not open to reasonable doubt.[69] But its philosophical foundations were given by Locke in his Essay. Essentially, its basis is the rejection of the infallible knowledge of objective necessary connections, the rejection of scientia. Once this goes the theory of legal evidence must be given a basis in empirical fact.

However, the point here is not to trace these developments in detail. Rather, it is to note that Bentham and others argue that a regress of probabilities tending towards zero occurs in a chain of testimony. We have just argued that this reasoning is, according to the calculus of probabilities, valid. But Hume's Sceptic's argument designed to extinguish all belief and evidence involves the same regress of probabilities. We must, therefore, accept the sceptic's argument as valid. Thus, contrary to critics like Prichard and MacNabb, Hume's argument for *sceoticism with regard to reason* is, if not sound, then at least very far from the "tissue of sophistries" that the received opinion charges it with being. Hume's

Sceptic's argument is, in short, an argument worthy of a philosopher, and, if it is unsound, then it is at least one that is worth refuting.

Fred Wilson
University of Toronto

1. Another case of the mind monitoring its own activities, and adjusting them to achieve its own ends, is the case of causal reasoning. See F. Wilson, "Hume's Theory of Mental Activity," in D.F. Norton et. al., McGill Hume Studies (San Diego: Austin Hill Press, 1983).

2. David Hume, A Treatise of Human Nature, ed. L.A. Selby-Bigge (London: Oxford University Press, 1888), p. 180.

3. Hume, (T182) speaks of a *reflex act of the mind:* feedback is a more modern term for the same thing.

4. Sextus Empiricus, Works, trans. R.G. Bury (Cambridge: Harvard University Press, Loeb Classical Library, 1935). Vol. II, pp. 331-3.

5. D. Hume, Enquiries concerning the Human Understanding and concerning the Principles of Morals, ed. L.A. Selby-Bigge (London: Oxford, 1902), p. 109ff.

6. EHV 111-13; T 181-2.

7. Cf. W. Salmon, Logic, First Edition (Englewood Cliffs, N.J.: Prentice-Hall, 1963), p. 63ff.

8. EHV 112. It has a physiological basis; cf. T 60-1.

9. T 311-73.

10. Treatise, I, III, XII. This reasoning is discussed in detail in F. Wilson, "Is There a Prussian Hume?" Hume Studies, 8 (1982), pp. 1-18. See also F. Wilson, "Mill on the Operation of Discovering and Proving General Propositions," Mill Newsletter, 17 (1982), pp. 1-14; and "Kuhn and Goodman: Revolutionary vs. Conservative Science," Philosophical Studies, forthcoming.

11. Ibid. This causal argument can be applied to the case of testimony, which involves causal connections (in Hume's sense) and also to the testimony of reason, for as we saw Hume remark (see fn. 2, above), *our reason must be consider'd as a kind of cause, of which truth is the natural effect.*

12. Cf. E. Nagel, Principles of the Theory of Probability (Chicago: University of Chicago Press, 1939), p. 44ff.

13. Cf. O. Sheynin, "Newton and the Classical Theory of Probability," Archives for the History of Exact Sciences, 7 (1970-1), pp. 237-8.

14. Cf. I. Lakatos, "Changes in the Problem of Inductive Logic," in I. Lakatos (ed.), The Problem of Inductive Logic, (Amsterdam: North Holland, 1968).

15. H.A. Prichard, "Hume," in his Knowledge and Perception (London: Oxford University Press, 1950), pp. 174-99.

16. Ibid., p. 174.

17. Cf. ibid., p. 189ff.

18. Cf. Wilson, "Hume's Theory of Mental Activity".

19. Cf. T. Beauchamp and A. Rosenberg, Hume and the Problem of Causation.

20. Prichard, Knowledge and Perception, pp. 195-6.

21. Ibid., p. 195.

22. Ibid.

23. Cf. D. Lewis, "Moore's Realism" Ch. V, in L. Addis and D. Lewis, Moore and Ryle: Two Ontologists (The Hague: Nijhoff, 1965).

24. Cf. Wilson, "Hume's Theory of Mental Activity." Compare also Hume's account of the causal relation between a volition and what the volition intends, in the Enquiry concerning the Human Understanding, Sec. viii.

25. D.G.C. MacNabb, Art. "Hume," in P. Edwards (ed.) Encyclopedia of Philosophy (New York: Macmillan, 1967), vol. IV, p. 84.

26. Compare T. Reid, Essays on the Intellectual Powers of Man, in his Works, ed., W. Hamilton (Edinburgh: MacLachlan and Stewart, 1852), Essay vii, Ch. IV, p. 487.

27. J. Bentham, Rationale of Judicial Evidence, in Works, ed. J. Bowring (London: 1838-43), Vol. 7, Bk. VI, Ch. III, p. 132.

28. Ibid., Ch. IV, p. 133.

29. G. Nokes, An Introduction to Evidence, Fourth Edition (London: Sweet and Maxwell, 1957), p. 23.

30. Cf. F. Wilson, "The Lockean Revolution in the Theory of Science," forthcoming in the festschrift for R.F. McRae; and "Critical Notice of I. Hacking's The Emergence of Probability," The Canadian Journal of Philosophy, 8 (1978), pp. 587-97.

31. For an important discussion of the scientia/opinio distinction, see I. Hacking, The Emergence of Probability (London: Cambridge University Press, 1975).

32. Cf. Hacking, The Emergence of Probability, pp. 20-4, p. 179.

33. Cf. Wilson, "Critical Notice of I. Hacking's The Emergence of Probability."

34. J. Locke, Essay concerning the Human Understanding, ed. A.C. Fraser (New York: Dover, 1959), vol. II, Bk. IV, Ch. XIV.

35. Ibid., Bk. IV, Ch. XIV, sec. 4.

36. Ibid., Bk. IV, Ch. XIV, sec. 6, p. 375.

37. Ibid., Bk. IV, Ch. XVI, secs. 9, 10; Bk. IV, Ch. XX, secs. 17, 18.

38. Ibid., Bk. IV, Ch. XVII, sec. 6, p. 401.

39. Ibid., Bk. IV, Ch. XV, sec. 6, p. 367.

40. Cf. Hume, Treatise, I, III, XII.

41. Just as the fact that the ideal gas law arises from extrapolating to the limit does not render that law is non-empirical; cf. F.W. Sears, An Introduction to Thermodynamics, Second Edition (Cambridge, Mass.,: Addison-Wesley, 1955), p. 7ff. J.L.

Mackie, *Truth, Probability and Paradox* (London: Oxford, 1973), p. 174, makes the same point.

42. G. Bergmann, "Frequencies, Probabilities and Positivism," *Philosophy and Phenomenological Research*, *6* (1945), p. 39.

43. Simon, Marquis de Laplace, *A Philosophical Essay on Probabilities*, trans. F. Truscott and F. Emory (New York: Dover, 1951), p. 4.

44. T I, III, XII, p. 132; I, III, XI.

45. The tossed coins are put into situations of unstable equilibria, where random small disturbances yield random upshots. But these unstable equilibria can be described in classical mechanical terms.

46. Cf. P. Martin-Lof, "The Literature on von Mises' Kollective Revisited," *Theoria*, *34* (1969), pp. 12-37. Also J.L. *Mackie, Truth, Probability and Paradox*, p. 173ff.

47. A. Kolmogorov, "On Tables of Random Numbers," *Sankhya: The Indian Journal of Statistics*, series A, *25* (1963), pp. 367-76.

48. A. Kolmogorov, Art. "Probability," in the *Great Soviet Encyclopedia*, trans. of the Third Edition, Moscow, 1970 (New York: Macmillan, 1974), vol. 4, pp. 423-4. L.E. Maistrov, *Probability Theory*, trans. S. Kotz (New York: Academic Press, 1974), is rather less judicious than is Kolmogorov when he too hastily condemns the work of von Mises.

49. Cf. Hacking, *The Emergence of Probability*, Ch. 11.

50. Cf. Bergmann, "Frequencies, Probabilities, and Positivism."

51. Cf. Nagel, *Principles of the Theory of Probability*, p. 34; and, with much more mathematical detail, W. Feller, *An Introduction to Probability Theory and Its Applications*, Third Edition (New York: Wiley, 1967), Ch. VI.

52. Cf. Hacking, *The Emergence of Probability*, p. 167.

53. Feller, *An Introduction to Probability Theory and its Applications*, Third Edition, pp. 150-3.

54. Cf. Hacking, *The Emergence of Probability*, pp. 167-8.

55. Ibid., p. 94.

56. Treatise, p. 132. Hume's talk of a "secret" operation of causes in this passage is of course compatible with his empiricism, and with operationism.

57. Cf. J.S. Mill, System of Logic, Eighth Edition (London: Longmans, 1961), p. 356, on how we judge the frequency of causes by the frequency of their effects. See also F. Wilson, "Goudge's Contribution to the Philosophy of Science," in L.W. Sumner, J.G. Slater and F. Wilson, eds., Pragmatism and Purpose (Toronto: University of Toronto Press, 1980).

58. The notion of "imperfect knowledge" is G. Bergmann's; see his Philosophy of Science (Madison: University of Wisconsin Press, 1957), Chapter II. The notion of "gappy" laws is J. Mackie's; see his "Causes and Conditions," American Philosophical Quarterly, 2 (1965). Compare J.S. Mill, System of Logic, p. 353, where the imperfect nature of statistical laws is recognized: we should not settle for chance when cause is possible.

59. I refer here, of course, to Hume's discussion *Of scepticism with regard to reason;* see T180, fn. 2 above.

60. Cf. the analysis of the argument from authority in W. Salmon's Logic. Salmon's text is one of the few to give a decent analysis of the logic of such arguments.

61. Cf. Nokes, An Introduction to Evidence, Fourth Edition, pp. 281-5. To put it in more detail: with hear-say evidence, the original giver of the evidence is not capable of being cross-examined in the courtroom, so that these standard procedures for discovering the presence of factors tending to bring about the assertion of falsehood are not available.

62. Nokes, ibid., p. 23.

63. J. Bentham, The Rationale of Judicial Evidence, ed. J.S. Mill.

64. Cf. ibid., Bk. I, Ch. VI, where Bentham proposes a scale on which witnesses might mark the degree of certainty of their testimony.

65. G. Bentham, Outline of a New System of Logi (London: 1827), p. 189ff.

66. System of Logic, p. 354.

67. Hacking, The Emergence of Probability, p. 178.

68. T. Waldman, "Origins of the Legal Doctrine o Reasonable Doubt," Journal of the History of ideas 20 (1959), p. 310.

69. Cf. Locke, Essay concerning Human Understanding Bk. IV, Ch. XV.