Belief and Beyond:

Toward a New Orientation in Epistemology

Prasanta S. Bandyopadhyay[1]

And

Mark L. Taper[2]

[1]Department of History and Philosophy, Montana State University,

Bozeman, MT 59717. <psb@montana.edu>

[2]Department of Ecology, Montana State University, Bozeman, MT 59717.

<taper@rapid.msu.montana.edu>

Abstract:

We draw a distinction between belief and evidence and develop two accounts to characterize these concepts. We argue that this distinction and the two subsequent accounts have significant consequences for traditional epistemology and the theory of inference by offering an alternative account of evidence and a diagnosis of the base-rate fallacy. We also discuss the old evidence problem in light of this distinction. In addition, we argue the superiority of our account of evidence by comparing it with that of an influential account. This paper calls for a new orientation in epistemology by providing a unified proposal for diverse philosophical issues. The belief/evidence distinction and the two accounts, we argue, are able to provide that orientation.

Comment: .

1. Overview
2. Belief, evidence, and their epistemological significance
3. Revisiting the belief/evidence distinction and its bearing on inference
4. Insights into some epistemological problems
    (i) The base-rate fallacy
    (ii) The old evidence problem
5. A comparison of this analysis with Achinstein's account of evidence
6. Beyond belief

1. Overview

The concept of belief lies at the heart of twentieth century epistemology. One possible reason for the preeminence of belief in epistemological discussions is that concepts like truth, falsity, and justification are attributed directly to beliefs and only indirectly to knowledge.[1] Belief is preeminent not only in different kinds of theories of justification but also in the literature on confirmation theories, especially in the standard Bayesian account, where Bayesians would prefer to talk in terms of an agent's degree of belief, rather than just the agent's belief.

Some contemporary epistemologists, however, emphasize other concepts like "evidence" as a central notion. Chisholm, for example, characterizes epistemology as "the theory of evidence."[2] However, there may be no more than a matter of emphasis at stake here. For example, Pollock and Cruz write,[3] "Epistemology might better be called 'doxastology,' which means the study of belief", while Chisholm calls it a theory of evidence. Both understand evidence in terms of belief and vice versa. Thus, in this prevailing view, if one has evidence for a proposition, one is justified in believing that it is true; and if one is justified in believing that a proposition is true, then one has evidence for it.

Many epistemologists overlook this distinction between belief and evidence. For example, after distinguishing knowledge from mere belief in the introduction to their anthology called *Knowledge,* Bernecker and Dretske, write, "[knowledge] is based on some form of justification, evidence, or supporting reasons."[4] Whereas justification is what an agent provides for believing a proposition; evidence for a proposition, as we will see, does not require believing that proposition. We will also see that under certain interpretations of evidence, it is possible that an agent's supporting reason be devoid of evidence. As a result, although each notion could be different from the other, Bernecker and Dretske are running all three notions together without any apology.

We believe that the last century's epistemological tradition has focused to an extreme on belief and has paid insufficient attention to that other key epistemological concept, evidence. The purpose of this paper is to challenge the tradition by arguing that the notion of evidence should be regarded as another central concept in epistemology along with the notion of belief. In order to lay out this challenge, we emphasize in sections 2 and 3 the ideas that belief and evidence are conceptually distinct notions and that conflating them leads to problems.

The fertility of a research program lies in its ability to resolve/respond to as many objections to the program as possible. Kahneman, Slovic and Tversky have raised an objection to Bayesianism for putting forward an "unrealistic" account of human rationality according to which an agent is rational if she follows the rules of the probability calculus and the expected utility theory.[5] They argue that the Bayesian account is unrealistic because human beings are not good at reasoning, probabilistic or otherwise. One way these critics explain human reasoning frailties is with the help of what is called the base-rate fallacy. Agents commit this fallacy in estimating the probability of an event when they overlook the base-rate associated, for example, with the prevalence of a disease in a population. They fail to factor that rate into their calculation when they calculate their own risk of having the disease. In section 4.1, we will discuss how the belief/evidence distinction advances an explanation of why agents overlook the base-rate the way they do

In addition to the charge that Bayesianism is unrealistic, Bayesianism is also confronted with a notorious problem known as the old evidence problem (OEP) posed by Glymour.[6] The problem is so called because Bayesianism allegedly fails to provide a satisfactory reconstruction of some of the episodes in the history of science in which old data have been used to support a new theory.[7] We argue that the OEP conflates the belief and evidence questions. A diagnosis of the problem reveals that the OEP has raised an evidence question, which our account of evidence is able to address. We will do this in section 4.2.

One often appreciates the strength of a new argument if one compares it with some existing account. Achinstein, for example, recently provided an account of evidence he thinks is adequate for the practice of science.[8] According to him, a necessary condition for a datum to be evidence for a theory is that it should provide good reason for believing the theory. By contrasting our account with his influential account, we highlight the superiority of our account especially because in associating evidence with belief, Achinstein conflates (strong) evidence and (justified) belief, thus failing to offer an account that is able to capture the practice of science. We discuss Achinstein's theory of evidence in section 5. We conclude by noting some advantages of our account in section 6.

2. Belief and Evidence: Epistemological Significance

Consider two hypotheses: H, representing that a patient suffers from tuberculosis, and ~ H, its denial. Assume that an X-ray, which is administered to the patient as a routine test, is scored as positive for the disease. Following the work of the statistician Richard Royall we pose three questions that expose the epistemological issue at stake in this simple scenario:[9]

(i) Given the datum, what should we *believe* and to what degree?[10]

(ii) What does the datum say regarding *evidence* for H and against its alternative ~H?

(iii) Given the datum what should we *do*?

The first question we call the *belief question,* the second the *evidence question,* and the third is the *decision question*. In this paper we primarily address the first two questions. Bayesians have developed two distinct accounts to address these questions; the first is an account of confirmation and the second one of evidence. For Bayesians, an account of confirmation provides a confirmation relation, C(D, H,B) among data, D, hypothesis, H, and the agents' background knowledge, B. Bayesians believe that because a confirmation relation is a belief relation in Bayesianism, it must satisfy the probability calculus, including the rule of conditional probability, as well as some reasonable constraints on one's *a priori* degree of belief in an empirical proposition. Learning from experience is, of course, part of Bayesianism. The rule of conditional probability ensures this inductive basis of our experience. As a result, the agent should not hold an *a priori* belief about an empirical proposition with full certainty (probability 1) or (probability 0) if she would like to learn from experience, but rather something in between these two extremes.

For Bayesians, then, belief is fine-grained. For an empirical proposition, they allow belief to admit of any degree of strength between 0 and 1. A satisfactory Bayesian account of confirmation should be capable of capturing this notion of degree of belief. In formal terms, it may be stated as:

> According to the confirmation condition (CC), D confirms H if and only if Prob(H|D) > 0 but < 1.[11]

The posterior probability of H could vary between 0 and 1 exclusive. Confirmation becomes strong or weak depending on how high or low this posterior probability is. The posterior and prior probabilities are essential parts of Bayes's theorem, a foundation of any Bayesian approach. According to the theorem, then, the posterior probability of a hypothesis H (Prob(H|D)) equals its prior probability (Prob(H)) multiplied by the likelihood of H (Prob(D|H)) then divided by the marginal probability of D (Prob(D))

$$\mathrm{Prob}(H \mid D) = \frac{\mathrm{Prob}(H) \times \mathrm{Prob}(D \mid H)}{\mathrm{Prob}(D)} \qquad \text{(EQ 1)}$$

The alternative evidential account lays down conditions for the evidence relation to hold among D, H, A as auxiliaries, and B. However, because the evidence relation is not a belief relation it need not satisfy the probability calculus. While this account of confirmation is concerned with belief in a single hypothesis embodied in equation 1, an account of evidence must compare the merits of two hypotheses, H1 and H2 (or ~H1), using data D.[12] Bayesians use the Bayes's factor (BF) to make this comparison, while others use the likelihood ratio (LR)[13] or other functions designed to measure evidence.[14] For simple hypotheses, the Bayes's factor and the likelihood ratio are identical[15] and capture the bare essentials of an account of evidence.[16] The BF/LR can be represented in this case by equation 2. [17]

D becomes E (evidence) for H1& A1& B against H2&A2 &B if and only if

$$\left[\frac{\Pr ob(D \mid H1,\ A1)}{\Pr ob(D \mid H2,\ A2)}\right] > k > 1 \qquad \text{(EQ 2)}$$

An immediate corollary of equation 2 is that there is equal support for both models only when BF/LR=1. Note that in (EQ2) k indicate a conventional threshold for the strength of evidence: if 1< K < 8, then D is often said to provide weak evidence for H1 against H2, while when K > 8 <16, D provides strong evidence, and when K > 16, D provides very strong evidence.[18] These cut-off points are determined contextually and may vary depending on the nature of the problem that confronts the investigator. The range of values for BF/LR varies from 0 to ∞ inclusive. Although the BF or LR behaves differently from a probability function in the sense that the former does not satisfy the rules of the probability calculus, there is a relationship between a likelihood function and a probability calculus: the former is proportional to the latter.

The crucial feature for understanding the distinction between an account of evidence and that of belief confirmation is the role of the “coherence condition.” Confirmation has a coherence condition (CCC) that says if H1 entails H2, then Prob(H2) ≥ Prob(H1). The coherence condition is a consequence of the axioms of the probability theory. As a result, any account of confirmation must satisfy it. However, evidence does not have a corresponding coherence condition (ECC) saying: If H1 implies H2, then whatever is evidence for H1 must also be (as good) evidence for H2.

A simple dice example will illustrate why the account of evidence does not need to satisfy this condition. Consider a pair of standard, fair dice, consisting of one red and one white die that have been thrown, but the result of the throw is unknown. The two hypotheses, H1 and H2, regarding the value of their faces would be:

(H1): the value on the faces of the two dice is equal to 3.

(H2): the value on the faces of the two dice is $\geq$ 3.

Note that H1 entails H2. Then an “evidence coherence condition” would imply that whatever is evidence for H1 must also be evidence for H2. Suppose our datum (D) is that we have learned that the red die showed 1. We have no information about the other die. Here, D favors H1 over H2, although H1 entails H2. The Bayes’s factor for H1 vs. H2 is $\frac{1/2}{5/35} = \frac{1/2}{1/7} = 3.5$. So although H1 entails H2, given D, whatever is evidence for H1 need not be as strong evidence for H2, thus violating the “evidence coherence condition (ECC).”

To explain the above scenario, one could say that D makes H2 less likely than it was before, while making H1 more likely. There were more ways H2 could be true than H1 before D is learned. However, once we know that D is the case, the occurrence of D then makes H1 a bit more likely than H2. Given D, H1 is more likely to be the case than H2, because there are so many ways that H2 could be true that are incompatible with D. It is important to realize that this dice example showing the violation of ECC does not anyway rely on any prior probability or belief for either of the hypothesis.

3. **Revisiting the belief/ evidence distinction and its bearing on inference**

Consider two diagnostic examples to show that there is a conceptual distinction between the belief and evidence questions. Let us first continue examining the tuberculosis example (TB) described earlier. The table summarizes a study estimating the efficacy of X-rays as a diagnostic tool for tuberculosis:[19]

X-ray	**Tuberculosis**		**Total**
	No	**Yes**	
Negative	1739	8	1747
Positive	51	22	73
Total	1790	30	1820

Table I.

Again, let H represent the hypothesis that an individual is suffering from tuberculosis and ~ H the hypothesis that she is not. These two hypotheses are mutually exclusive and jointly exhaustive. In addition, assume D represents a positive X-ray test result. Let us solve for Prob(H|D), the posterior probability that an individual who tests positive for tuberculosis actually has the disease. Bayes's theorem helps to obtain that probability.

To use the theorem, we need to know first Prob(H), Prob(~H), Prob(D|H), and Prob(D|~H). Prob(H) is the prior probability that an individual in the general population has tuberculosis. Because the 1820 individuals in the study shown in Table I were not chosen from the population at random, a valid frequency based prior probability of the hypothesis couldn't be obtained from it. Yet in a 1987 survey, there were 9.3 cases of tuberculosis per 100,000 people in the general population (Pagano and Gauvreau, 2000). Consequently, Prob(H) = 0.000093. Hence, Prob(~H) = 0.999907. Based on Table I, we may compute the following probabilities. Prob(D|H) is the probability of a positive X-ray given that an individual has tuberculosis. Prob(D|H) = 0.7333. Prob(D|~H), the probability of a positive X-ray given that a person does not have tuberculosis, is = 1 − Prob(~D| ~H) = 1− 0.9715. = 0.0285.

Using all this information, we may compute Prob(H|D) = 0.00239. For every 100,000 positive X-rays, only 239 signal true cases of tuberculosis. The posterior probability is very low, although it is slightly higher than the prior probability. Although the confirmation condition (CC) is satisfied, the hypothesis is not very well confirmed. Yet at the same time, the BF, i.e., 0.7333/0.0285 (i.e., Prob (D|H)/Prob(D|~H)) = 25.7, is very high. Therefore, the test for tuberculosis has a great deal of evidential significance.

Second a case where an HIV specialist wants to know whether a prostitute chosen at random is afflicted with the virus. He administers the only available test. H is the hypothesis that the prostitute has the virus. Given his expertise in the field(and without firm relative frequencies to guide him), he has assigned 0.99 in advance to the hypothesis that the prostitute is carrying the HIV virus. D represents the positive test outcomes, ~D means that test turns out to be negative. In this example, there are two questions the specialist wants answered. One concerns the probability that the prostitute in question is afflicted with the virus given that the test says that she is, i.e., what's Prob(H|D)? The other has to do with the quality of the test, with what we might call its evidential significance.

To answer these questions, he needs to know the likelihoods of the data under H, and ~ H respectively. To know these, the specialist needs to run some tests. Suppose these result in the following likelihoods. Prob(D|H) = 0.1625, Prob(~ D| H) = 0.8375, Prob(D|~H) = 0.8136, and Prob(~ D|~ H) = 0.1864. One could see from these likelihoods that the tests are unreliable. We compute Prob(H|D) = 0.94, a small decrease over the specialist's initial degree of belief. To this extent, the hypothesis is less confirmed; he has slightly less reason now to believe that this prostitute has HIV. However, the test provides no evidence for the hypothesis as against its denial. Prob(D|H)/Prob(D|~H) is 0.199. Which is to say that significance of the evidence is much less than the degree to which the original hypothesis has been, in the sense indicated, "disconfirmed." Indeed, the ratio provides evidence against H in favor of ~H.

There are cases, then, in which the evidence is good, but the hypothesis is minimally confirmed, and in which the evidence is poor, but the hypothesis receives a significant degree of confirmation. Evidence and confirmation are to be distinguished with respect to what we have called their "contextual" features, the former presupposing and the latter not a comparison between hypotheses. They are to be distinguished with respect to the fact that confirmation is and evidence is not a probability measure. Finally, there is no linear relationship between them; confirmation and evidence in particular cases vary independently, small degree of confirmation and good evidence in some cases, high degree of confirmation and poor evidence in others. All of this might have expected given that confirmation has to do with belief and its degrees over time, evidence with a relationship between hypotheses considered jointly and certain kinds of data.

There is little point in denying that the meanings of "evidence" and "confirmation" (or its equivalents) often overlap in ordinary English as well as among epistemologists. There is a theorem, BF > 1 if and only if Prob(H|D) > Prob(H), which shows this connection.[20] However, strong belief does not necessarily imply strong evidence, and strong evidence does not necessarily imply strong belief, as illustrated by the TB example. Our case for distinguishing them rests not on usage, but on the clarification thus achieved in our thinking and inferences frequently made in diagnostic studies.

4. Insights into some epistemological problems

4.1 The base-rate fallacy

Uncertainty is an inescapable aspect of human life. Many of our significant choices are based on partial information about uncertain events. More recently, Kahneman, Slovic and Tversky have argued that, when in an uncertain world we are confronted with choices regarding what is likely to happen, we are more often than not far from being perfect reasoners. They demonstrate case after case in which human beings are bad in this type of probabilistic reasoning. To illustrate how human agents could make probabilistic errors, they provide the base-rate fallacy as an example. One way to discuss the fallacy is to use a variant of the TB example where the agent is posed with the following problem, along with a question at the end.[21]

The agent is worried that she has tuberculosis. One percent of the population has tuberculosis. There is a simple and effective test, which identifies tuberculosis in everyone who has it, and only gives a false positive result for 10 percent of people without it. She takes the test, and gets a positive result. Which of the following is closest to the probability she has the tuberculosis?

(A) 90 percent

(B) 9 percent

(C) 50 percent and

(D) 89 percent.

Typically the agent fails to give the correct response. The average estimate of the probability that she suffers from the fictional disease was 85 percent, whereas the correct answer based on Bayes's theorem is 10 percent. Why does the agent typically fail in making the correct choice here? According to Kahneman, Slovic and Tversky, she fails because she overlooks the base-rate of the disease in the population -- that is, only 1 percent of the population suffers from the disease.

Their understanding of these psychological data has come to be known as the "heuristics and biases" approach, which explains the data by arguing that human beings adopt certain heuristic strategies in solving theoretical problems, strategies which often provide useful short-cuts to reasonably accurate answers, although these strategies can be shown to bias subjects irrationally toward certain kinds of mistakes.

However, we argue that the belief/evidence distinction is able to provide an alternative rationale for the agent's failure. Specifically, we argue that in the case of the base-rate fallacy the agent (not the investigator) is conflating the belief and evidence questions. The former thought that the investigator has asked the evidence question, whereas the latter is really asking the belief question. Because the subject thought so based on the data about true positive and false positive of the X-ray, she took the data to provide strong evidence for the presence of tuberculosis as against not having the disease. Given the data, the Bayes's factor/ likelihood ratio based account of evidence (i.e., Prob(D|H)/ Prob(D|~H) = 1/0.1= 10) demonstrates that under this data having the disease is 10 times more likely than not having the disease, thus providing support for her intuitive judgment about the strength of the disease. This is *strong evidence* for the hypothesis that she is likely to have tuberculosis relative to not having the disease. So the reason that the agent overlooks the base-rate is that she mistakenly takes the question, "what is now the probability that she has tuberculosis?" to be an evidence question. As a result, given the positive test result, she thinks that it is very likely that she has it, whereas the probability of her having the disease is, in fact, just one tenth.[22]

4.2 The old evidence problem

Perhaps the most celebrated case in the history of science in which old data have been used to construct and vindicate a new theory is that of Einstein. He used the

Mercury's perihelion shift (M) to verify the general relativity theory (GTR).[23] The derivation of M has been considered the strongest classical test for GTR. However, Einstein was aware of the shift before he constructed his theory.[24] So this is a case of the old evidence problem (OEP)[25]. According to Glymour, Bayesianism fails to explain why M is regarded as evidence for GTR. For Einstein, Prob(M) = 1 because M was known to be an anomaly for Newton's theory long before GTR came into being. But, Einstein derived M from GTR; therefore, Prob(M|GTR) = 1. Glymour contends that given (EQ 1), the conditional probability of GTR given M is therefore the same as the prior probability of GTR; hence, M cannot constitute evidence for GTR.

Consider why no account of confirmation could be of help for the OEP. Regarding Mercury's perihelion shift (M) and GTR, Prob(M) is 1 (or nearly so) only because Prob(M|M) is 1, which is a conditional probability. However, within the framework of the OEP, Prob(M) is the marginal probability. Thus:

$$\Pr ob(M) = \Pr ob(M \mid GTR) \times \Pr ob(GTR) + \Pr ob(M \mid Newton) \times \Pr ob(Newton) + \Pr ob(M \mid Catch-all) \times \Pr ob(Catch-all) \qquad \text{(EQ 3)}$$

In (EQ 3) "Newton" represents Newton's theory of gravitation and the "Catch-all" stands for those logically possible theories, which are yet to be conceived, but are taken to be different from both GTR and Newton's theory. If there is a universal theme in science, it is that all theories are eventually replaced or refined. One could argue that our best known theories including the GTR are likely to be wrong theories in the end. One motivating reason for it is that even GTR is unable to unify relativity theory with quantum mechanics. Thus, we could have a strong belief that there are such theories even in the presence of the GTR. Consequently, the catch-all must be included in equation 3 to capture this belief. However, neither Prob(M|Catch-all) nor Prob(Catch-all) can be assigned values, therefore equation 3 can no longer be evaluated. So we can't have any confirmation for GTR.

For the sake of argument assume assignments of non-zero prior probability for the Catch-all and a non-zero value of the likelihood of M under the Catch-all. Furthermore it is known that Prob(M|GTR) = 1 and Prob(M|Newton) = 0. Given this information about the Catch-all hypothesis and the likelihoods of Newton's theory and GTR, it follows from equation 3 that Prob(M) must be > Prob(GTR), when Prob(M) = 1. What further follows from this result for the old evidence problem is that for M to be evidence for GTR, M must be believed or known to be true. So the assumption on which the old evidence problem rests is that unless the agent believes or knows M to be true, M could not be counted as evidence for GTR.

In addition, for the marginal probability of M to be equal to the right hand side of EQ3, Prob(GTR) has to be very nearly one when both the prior and likelihood of the Catch-all have nonzero (but very close to zero) values. The GTR to have a prior probability close to 1 entails that virtually the agent is committed to the truth of GTR. Before the theory has been tested, if a theory has a probability close to 1, then the theory has been virtually confirmed. Moreover, this probability will not change at all as a function of either new or old data. This consequence is unwelcome for any account of confirmation.

To sum up: there are three reasons why one can't expect a confirmation theory to address the old evidence problem. First, because we are unable to compute the likelihood of the Catch-all hypothesis, we would have no way to compute the value in EQ 3. Second, to assign probability 1 to M, which is an empirical proposition, forces the agent to ignore experience. Learning from experience is a crucial aspect of any reasonable epistemology including Bayesianism. Finally, to be able to assign probability 1 to M entails that the agent in question is forced to assign prior probability of almost 1 to GTR before the theory has been tested. Based on theses considerations, we conclude that although in the history of science it is well-know that M provides strong support for GTR, an account of confirmation, however, fails to deliver that needed support for GTR.

We cite the passage where Glymour has discussed the problem, and for the sake of clarity divide it into A and B.

(A) "The conditional probability of T on e is therefore the same as the prior probability of T,….

(B) "e cannot constitute evidence for T,…." [26]

Concerning (A), we take it that Glymour considers conditional and prior probabilities as concerned with agents' degrees of *belief*, and conditional probability prescribes a rule for *belief revision*. On the other hand, (B) states that e can't be regarded as *evidence* for T.

We argue that (A) pertains to the belief question, while (B) pertains to the evidence question, and that separate accounts are needed to address them. Glymour conflates (A) and (B). Separate accounts of confirmation and evidence can be provided to handle each, as outlined above. However, as a result, although the OEP is a problem about evidence and hence needs an account of evidence to tackle it, it is mistakenly taken to be asking a belief question as well, and therefore looking for an account of confirmation for its rescue. We argued that confirmation theory can't account for the historical success of GTR by Mercury's shift.

An account of evidence based on the Bayes's factor/the likelihood ratio is, however, able to address (B) that pertains to the evidence question. Consider GTR and Newton's theory relative to M with different auxiliary assumptions for respective theories. Two reasonable background assumptions for GTR are (i) the mass of the earth is small in comparison with that of the sun so that the earth can be treated as a test body in the sun's gravitational field and, (ii) the effects of the other planets on the earth's orbit is negligible. Let A1 represent those assumptions.

For Newton, on the other hand, the auxiliary assumption is that there are no masses other than the known planets that could account for the perihelion shift. Let A2 stand for Newton's assumption. When the Bayes's factor is applied to this situation, there results a factor of infinity of evidential support for GTR against its rival. For, $Prob(M|GTR \& A1\&B) \approx 1$, where $Prob(M|Newton \& A2 \& B) \approx 0$. Hence, the ratio of these two expressions goes to infinity as Prob(M|Newton &A2 &B) goes to zero. The likelihood function does not satisfy the rules of the probability calculus. Plus, its evidential support of being close to infinity is consistent with the notion of likelihood being a well-defined function. So when the old evidence problem is correctly diagnosed as asking an evidence question, there is a satisfactory account of evidence to address it.

One possible worry among some Bayesians is whether it is at all possible to discuss the notion of evidence without falling back on (high) prior probability or belief in the theory which has been declared the winner.[27] If this worry were justified, then it should also cut across any account of evidence including our BF/LR based account, which is intended to provide a comparative account of two simple hypotheses without falling back on their prior probabilities. We will provide two responses to this worry. First, we argue that it would be irrelevant to talk about prior probability (high or low) or belief when one is concerned with the evidence question. Second, for the sake of argument even though we concede that above worry has some plausibility that it won't be irrelevant to talk about prior probability when one deals with the evidence question, we will contend that the worry is *not* generally true. That is to say, even though the BF/LR based account yields high value for a theory, when the probability of the data is very high, it does not follow from it that the prior probability for that theory must be then very high.

Consider the first response that when one is concerned with the evidence question, it would be irrelevant to ask for the prior probability of the theory in question. To understand our response, let's recall the HIV case from section 3. In that example, the specialists would like to have two questions answered. They are (i) the belief question, that is, what should we believe and to what degree about the hypothesis that the prostitute in question is suffering from the HIV virus?, and (ii) the evidence question, that is, given the test, what does the test say about evidence for the hypothesis as against its alternative? Given the belief question, we have seen that the specialist finds that he should believe strongly in the hypothesis, although the latter's posterior value is slightly lower than its prior value. Assume that a second authority suggests that the specialist in our example should not believe the hypothesis that the prostitute is suffering from the HIV, because the evidence goes against the hypothesis in favor of its alternative. Under this scenario, the specialist in our story could respond that the second specialist is giving an irrelevant response. For, while the former is asking a belief question in which the high confirmation value is relevant, the second specialist recommends the former to use the BF/LR based value appropriate only for an evidence question. In short, if one is asking a belief question, then it is irrelevant to provide an evidence-based response. In the same vein, it would be irrelevant to contend that a high BF/LR value for a theory entails a high prior probability for that theory when the probability of the data is very high. Otherwise, it would be to muddle two epistemological questions, the evidence question and the belief question. Since, the old evidence problem has asked an evidence question and an account of evidence is able to handle it, it follows that it would be irrelevant to observe that "we are already committed to believing strongly that theory for which there is strong evidence, when the probability of the data is very high". [28] This observation would be irrelevant in the same way if the second specialist in the HIV case would say that the first specialist should not believe that the prostitute is suffering from the disease because evidence goes against the hypothesis.

Second, for the sake of argument, let's first assume that above response is incorrect, that is, one could apply tools for confirmation theory to respond to the evidence question. Despite this assumption, we will argue that our BF/LR based resolution of the old evidence problem does not *necessarily* rest on making the winning theory to have a high prior probability of being true. The moral of the second response will be that the objection -- if both the BF/LR is very high and the probability of D is very high being an old evidence, then the winning theory will be already assigned a very high probability-- is not generally true. Consider the argument for our claim. (EQ 3) is reduced to (E4) since Prob(M|GTR) = 1 and Prob(M|Newton) = 0.

$$\Pr ob(M) = \Pr ob(GTR) + \Pr ob(M \mid Catch-all) \times \Pr ob(Catch-all) \qquad \text{(EQ4)}$$

Equation 4 implies that Prob(GTR) = Prob(M) − Prob(Catch-all & M). Although we can't evaluate Prob(Catch-all & M), we can easily see that Prob(GTR) is not constrained to be near one. It could be anything between 0 and Prob(M) depending on the value of Prob(Catch-all & M).[29] In the case of our account, we investigate the evidential relation between two theories independent of whatever the agent might believe about those theories. It is because in the case of the old evidence problem, we are interested in the evidence question to explain why Mercury's shift provides support for GTR. One point to note is that our resolution of the old evidence problem is general and is not dependent on a specific case study like the classical support of GTR by Mercury's shift. This resolution should be applicable to all sorts of the so-called old evidence problem for both deterministic theories like GTR and statistical theories.

Two comments are in order. The first concerns the criterion of evidence on which the account of the old evidence problem implicitly rests. The criterion of evidence Glymour exploits to generate the old evidence problem is that D is evidence for H if and only if Prob(H|B&D) > Prob(H|B). According to the old evidence problem, the agent's belief in H should increase when the criterion holds. This makes clear how the issue of what the agent should believe is embedded in the criterion for evidence.

The second comment concerns the implication of Glymour's evidence criterion regarding the relationship between evidence, knowledge, and belief. According to Glymour's OEP, whenever the criterion of evidence is satisfied, D becomes evidence for H relative to an agent. D being evidence for H does not, however, turn out to be independent of whether the agent knows or believes D to be the case. By contrast, in our account whether D is evidence for H1 vs H2 is independent of whether the agent believes or knows D.[30]

5. A comparison of this analysis with Achinstein's account of evidence

Achinstein begins his book with what he calls the "dean's challenge" regarding why the dean of his university thinks that philosophers of science have made almost no contribution to scientists' understanding of evidence, explanation, or laws. To provide a response to the dean, he investigates closely the notion of and problems associated with evidence. Although he distinguishes various types of evidence, for him, veridical evidence is that which is alone able to capture the scientist's notion of evidence. He argues that "if e is evidence that h [i.e., hypothesis], then" veridical evidence implies that "e provides a good reason to believe h." [31] Call this condition on veridical evidence C1. He explains the consequent of the above conditional statement in the following way: "[f]or any e and h, if e is a good reason to believe h, then e cannot be a good reason to believe the negation of h (~h)." [32] This condition, he thinks, provides necessary, but not sufficient condition for reasonable belief. Spelling out the last condition along with some further constraints on evidence, he provides what we call C2, which says, "e is a good reason to believe h only if Prob(h|e) > 1/2". [33] These two conditions characterize his account of veridical evidence.

Achinstein thinks that scientists seek only veridical evidence because they would like to have a good reason for believing in their theories, in the sense that "good reason" requires truth. He argues further that requiring probability of a theory, given the data greater than one half when an agent has good reason for believing a theory, provides one aspect of his response to the dean, because this imposes a stronger requirement on an account of evidence than requiring its probability be greater than zero.

Consider again the TB example of section 3. If we apply Achinstein's C1 to the example, then it follows that, since the data provide strong evidence for the presence of tuberculosis, the agent should believe that she has tuberculosis. However, that evidence does not necessarily provide good reason to believe that the person in question is more likely to have the disease than its absence. The posterior probability of H (i.e., 0.00239) is slightly higher than its prior probability (i.e., 0.000093). So contra Achinstein, even though e provides strong evidence for H, e does not provide a good reason for believing H. The posterior probability of the denial of the hypothesis is 0.9976, which is considerably higher than one half. In contrast, if we consider C2, then we are forced to admit that we have good reason not to believe that the person is suffering from TB because the denial of the hypothesis is higher than one half, whereas C1 implies that, since we have (strong) evidence for the hypothesis that the person in question is suffering from tuberculosis, then we should have good reason to believe the latter. In the TB case, Achinstein's account of evidence generates a tension regarding what one should really believe. Should we believe the hypothesis that the person is likely to have tuberculosis because the evidence says so, or should we rather disbelieve it because the latter's posterior probability is lower than one half?

On deeper scrutiny, the tension in Achistein's account may be traced back to his failure to distinguish adequately between strong evidence and justified belief. He takes evidence to be something that should provide good reason for belief. However, we have argued that belief and evidence are conceptually distinct concepts. Although there is a relationship between them, the reason that his account has generated a tension is that while evidence could strongly favor the hypothesis that the person has the disease, the agent's belief could still go in the opposite direction. The discussion of the TB example just offered illustrates this divergence. Because Achinstein's account fails to distinguish adequately between strong evidence and justified belief, it generates a counter-intuitive consequence, thus yielding a methodology inadequate to address his own dean's challenge.

There are additional shortcomings in Achinstein's account. The first has to do with the conflict between his quantitative threshold of ½ given in C2 and his qualitative condition of "good reason to believe". The second is that his account is not a comparative one. Consider the first shortcoming. He fails to realize that what counts as evidence need not count as "good" evidence. Suppose a coin is tossed five times and comes up with heads four times. We consider two models: a "biased coin model" with the probability of heads equal to 0.8, and a "fair coin model" with the probability of heads equal to 0.5. The sequence of four heads and one tail is certainly evidence for the biased coin model relative to the fair coin (LR(biased/fair) = 2.62). [34] However, it is weak evidence because there is a reasonable chance that a fair coin might also give such a sequence. Hence that sequence of throws does not constitute a "good reason to believe" that the coin is biased.

Although the cutoffs are arbitrary, the distinction between weak and strong evidence allows both for evidence to be continuous, and for some strength of evidence, whether strong, or very strong, to constitute a "good reason to believe" to use Achinstein's expression. Another reason why it is important to consider weak evidence is that while a single small piece of evidence by itself need not be compelling, many small pieces of weak evidence could be accumulated into strong evidence. Consider, a scenario in which four more people each toss this coin five more times and each ends up with the same result of four heads. The evidence for a biased coin in each person's sequence of tosses is weak. However, the combination of their evidence is compelling (LR(biased/fair)=123.4).[35] Based on these considerations, we have argued that Achinstein's account of evidence fails to take into account these nuances of evidence that are at the heart of the practice of science Achinstein is so keen on capturing.

A further defect in Achinstein's account is that his measure of evidence does not compare hypotheses. Consider three examples involving real data sets that investigate whether there is a positive relationship between smoking (S) and aortic stenosis (A). The latter is a disease that represents narrowing of the aorta that prevents the flow of blood to the body.[36] In the first example (Table II), investigators are interested in finding whether there is an observed relationship between S and A among males. So is the interest in the second example where the only difference is that this time it is confined to females (Table III.) In the third and final one, investigators will study the same relationship in the amalgamated data (Table IV) taking into account both data sets from the preceding Tables (II and III).

Males

Aortic Stenosis	**Smoker**		**Total**
	Yes	**No**	
Yes	37	25	62
No	24	20	44
Total	61	45	106

Table II

Females

Aortic Stenosis	**Smoker**		**Total**
	Yes	**No**	
Yes	14	29	43
No	19	47	66
Total	33	76	109

Table III

If we compute the conditional probabilities of A given S [i.e., Prob(A|S)], and of A given not S [i.e., Prob(A|~S)] as recommended by Achinstein, then we obtain Prob(A|S) = 37/61 = 0.60 and Prob(A|~S) = 0.55 based on Table II. Since according to him, both probabilities are greater than ½, we can't rule out these values as not being evidence for the presence of the disease. In his account, those values only satisfy a necessary condition for some data to be evidence. In contrast, if we calculate those same conditional probabilities for females based on **Table** III, we get Prob(A|S) = 14/33 = 0.42 and Prob(A|~S) = 0.38. Because these probabilities for females are lower than 1/2, Achinstein is committed to saying that there is no evidence for a positive relation between aortic stenosis and smoking.

However, clearly there is an observed relationship between aortic stenosis and smoking in both data sets. It is with the help of a comparative account of evidence alone, that one could appreciate the relationship between those two variables independent of the influence of gender. Consider our account of evidence based on the BF/LR. If we apply our account to two data sets, then we will obtain exactly same values for both males and females we have obtained applying conditional probabilities without committing the "transposing the conditional fallacy.[37] We are able to identify two conditionals together and the latter would yield the same value only in that special case in which the prior probability for A and ~ A to be equally probable in both populations. Since this special condition is met, these transposed conditionals, for example, Prob(A|S) and Prob(S|A), imply the same value. We find that the LR for S vs. ~S conditional on A = 0.60/0.55 = 1.09 relative to the male population, and LR for S vs. ~S conditional on A = 0.42/0.38 = 1.10 relative to the female population. Independent of gender, there is an influence of

Comment:

Comment:

smoking on aortic stenosis, which only a comparative account of evidence could bring out. The relationship we have detected in both the data sets is that there is a greater chance of having aortic stenosis among smokers than among non-smokers. An analysis comparing hypotheses exposes this relationship while an account based on confirming single propositions fails to do so.

If one amalgamates both the data from both Tables II and III and form Table IV, then one realizes that the amalgamated table manifests a case of Simpson's paradox.

The Amalgamated Data

Aortic Stenosis	**Smoker**		**Total**
	Yes	**No**	
Yes	51	54	105
No	43	67	110
Total	94	121	215

Table IV

This is a version of Simpson's paradox in which the magnitude of the relation of two variables (smoking and aortic stenonis) is influenced by the presence of a third variable (gender). If the effect of gender is disregarded, then the strength of the association between smoking and aortic stenosis appears greater than it is for either the male population or female population alone.

Simpson's paradox is not a paradox by itself because it rests on simple arithmetical relations among numbers in the sense that if the data in Tables II and III are correct, then the data set Table IV displays must follow necessarily. Since Simpson's paradox is not a paradox, if an account generates this so-called paradox, the account should not thereby be regarded as faulty. However, Achinstein's account of evidence is faulty, because it fails to provide a satisfactory account of the emergence of Simpson's paradox displayed in the relationship among three categorical variables, (see Table II, III and IV). Again, this failure on the part of his account is traceable to his account not being comparative.

If we look at the amalgamated data (Table IV), his account can't rule out Prob(A|~S) = 0.51 as evidence for a positive relationship between smoking and aortic stenosis, because the amalgamated data only satisfy the necessary condition of his account of evidence, although the data provide evidence for a positive relation between smoking and aortic stenosis. The BF/LR based account brings out this positive relation between smoking and aortic stenosis for the amalgamated data correctly. In the latter, the BF/LR value is $\frac{51/94}{54/121}$ = 1.23 than the BF/LR value for the same relationship among males, which is 1.10 and among females, which is, 1.09 respectively. Because we have a comparative account, we could account for the generation of Simpson's paradox by comparing respective BF/LR values in three tables, (Table II, III and IV) in terms of their different ratios in Tables II and III as compared to Table IV. Simpson's paradox is a thorny problem frequently besetting real scientific analysis. Any account of evidence, including that of Achinstein, that fails to explain why Simpson's paradox arises is inadequate to buttress statistical inferences at the core of the scientific practice.

6. Beyond belief:[38]

Most epistemologists so far have paid little or no attention to the concept of evidence, treating it either as definable in terms of belief or subordinate it to the latter. This paper challenges the tradition by providing a new orientation in terms of a distinction between the belief and evidence questions. It explains why the concept of evidence holds a key to several epistemological issues and problems in theory of inference. The notion of evidence this paper advocates makes the status of something as evidence independent of whether it is something that the agent knows or believes. The data D is evidence for H (or not), independent of whether the agent happens to believe D, or indeed independently of whether it is evidence the agent has any sense.

One question that has often been posed against such an account of evidence is that "what is the value for such an account if it does not provide a good reason, or indeed any reason, for believing that which it is evidence for?" or as Achinstein puts it "why is this a concept of *support* or *evidence*?" (p.130). First, we will argue that the question rests on an assumption that evidence could not be misleading. Then, we will raise further problems with this objection.

Consider the first assumption on which the question rests, that is, evidence could not be misleading. Take our coin toss example with the two models: the "biased coin model" with the probability of heads equal to 0.8, and the "fair coin model" with the probability of heads equal to 0.5. The first sequence of coin tosses with four heads and a tail was taken as evidence for the biased coin model over the fair coin model. Subsequent data strengthened the evidence. But what if the following sequences had not all yielded 4 heads but had instead had 2, 3, 2, and 1 heads. With 12 heads out 25 tosses, the likelihood ratio for the biased coin model relative to the fair model is 0.00189. This is not even weak evidence for the biased coin model. It is in fact very strong evidence for the fair coin model over its counter-part. Were we wrong to take the first 5 tosses as evidence for a biased coin? The answer is “no”, because the evidence was weak. We have been foolish to assign too much credence to it. It is crucial to realize that evidence can be misleading.[39] However, as the evidence becomes strong, fortunately the probability of misleading evidence becomes low.[40] Thus the important lesson to learn from this discussion is that the link between evidence and belief need not be as tight as proposed by Achinstein and other epistemologists, because the evidence on which belief might rest can be misleading.

Once it is possible to have evidence without belief, evidence become transferable. Two agents with differing background information and beliefs can not readily exchange information using confirmation based measures such as posterior probabilities. Information can be easily and compactly exchanged as evidence. In effect, the evidence summarizes what the data say about the compared models. Each agent can subsequently consider their own beliefs, background knowledge and auxiliary information in reaching a conclusion.

Achinstein asks the question what’s the value of such an account of evidence and what is evidence good for? To answer this, we need to turn from our rather esoteric considerations to a very pragmatic understanding of what evidence is in the real world. A common language definition of evidence is that it is “a thing or things helpful in forming a conclusion or judgment.”[41] So evidence, while not compelling us to belief can contribute to our beliefs or decisions, which may also be influenced by other pieces of evidence or belief.

For example, consider a woman who has just taken a PAP test that screens whether she is suffering from cervical cancer. Assume that the test yields a positive result, and the test is quite reliable, although it is not perfect. Given the positive result for the PAP test, the agent could ask all three types of questions, the belief question, the evidence question, and the decision question. For the sake of argument, we could assume further that the evidence gleaned from her positive PAP result is weak for the presence of cervical cancer in the patient. However, let's assume that the frequency-based prior for an individual having the disease is very small. Now what she should do? Here, evidence along with the belief that rests on the frequency-based prior for an individual to have cervical cancer might help her make a decision about what she should do. She might factor into her age in her decision whether to opt for a biopsy or see another doctor for further input into her decision. Thus, evidence could help the patient to make a decision about what she should do regarding her PAP test.

We emphasize that we don't suggest that evidence could never provide reasons for believing a theory or a proposition. There are clearly cases in which we have both strong evidence and strong belief in a hypothesis[42]. In the latter case, we could say that we are justified in believing that hypothesis. However, how to combine these features, strong evidence and strong belief, to develop a full-blown theory of justification and knowledge that is able to address several notorious epistemological puzzles like the Gettier problem, the Grue paradox and the lottery paradox remain a desideratum for further investigation.

Select Bibliography

Achinstein, P. (2001): *The Book of Evidence*. Oxford, Oxford University Press.

Achinstein, P. (1983): *The Concept of Evidence*. (ed.)Oxford, Oxford University Press

Alston, W. (1989): *Epistemic Justification*. Ithaca, Cornell University Press.

Armstrong, D. (1973): *Belief, Truth and Knowledge*. New York, Cambridge University Press.

Audi, R. (1993): *The Structure of Justification*. New York, Cambridge.

Bandyopadhyay. P. S. and Brittan. G. Jr. 2005. Acceptability, Evidence, and Severity. *Synthese*, p.1-36, Feb.

Berger, J. O. (1985) *Statistical Decision Theory and Bayesian Analysis*. New York, Springer-Verlag.

Bernecker, S. and F. Dretske. (2000): *Knowledge*. (eds.) Oxford, Oxford University Press.

Bonjour, L. (1985): *The Structure of Empirical Knowledge*: Cambridge, Harvard University Press.

Bonjour, L. (1992): "Externalism/Internalism" in Dancy and Sosa (eds.): *A Companion to Epistemology*. Oxford, Blackwell.

Bovens, L. and S. Hartmann. (2003): *Bayesian Epistemology*. Oxford, Clarendon Press.

Box, G. E. P. (1976) Science and Statistics. *Journal of the American Statistical Association*: 71(356) pp 791-799.

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

Chisholm, R. (1977): *Theory of Knowledge*. Upper Saddle River, Prentice Hall (second edition).

Chisholm, R. (1980): “A Version of Foundationalism” in P. French, H. Wettstein, and T. Uehling, ed., *Midwest Studies in Philosophy 5: Studies in Epistemology*. Minnesota, University of Minnesota Press.

Christensen, D. (1999): “Measuring Confirmation.” *Journal of Philosophy*, 99, No 9, Sept, pp.437-461.

Dennett, D. (1987): “Beyond Belief.” *The Intentional Stance*. Cambridge. MA: MIT Press.

Earman, J.(ed.) (1983): *Testing Scientific Theories*. Minneapolis, University of Minnesota Press.

Earman, J. (1992): *Bayes or Bust?* Cambridge, MA: MIT Press.

Earman, J and Salmon, W. (1992):

Feldman, R. and E. Conee. (2001): “Internalism Defended.” in *American Philosophical Quarterly*, Vol 38, No 1, Jan, pp.1-17.

Fitelson, B. (1999) The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity. *Philosophy of Science*, Supplement to vol 66, no. 3 pp 362-378.

Gill, J. (2002): *Bayesian Methods*, London. Chapman & Hall.

Glymour (1980) *Theory and Evidence*. Princeton. Princeton University Press.

Goldman, A. (1986): *Epistemology and Cognition*. Cambridge, Harvard University Press.

Goldman, A. (2002): *Pathways to Knowledge*. Oxford, Oxford University Press.

Good, J. (1983): "Some Logic and History of Hypothesis Testing." in *Good Thinking* Minneapolis, University of Minnesota Press.

Greco, J. and E. Sosa. (1999): *The Blackwell Guide to Epistemology*. (Eds.): Malden, Mass.: Blackwell.

Han, J. and M. Kamber. (2001): *Data Mining*. San Francisco: Morgan Kaufmann.

Heil, J. (1992): "Belief." in Dancy and Sosa (eds.): *A Companion to Epistemology*. Oxford, Blackwell.

Howson, C. (2000): *Hume's Problem*: Clarendon Press: Oxford.

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

Joyce, J. (1999): *The Foundations of Causal Decision Theory*. Cambridge, Cambridge University Press.

Kahneman, D., Slovic, P., and Tversky, A. (1982): *Judgment under Uncertainty*: *Heuristics and Biases*. Cambridge: Cambridge University Press.

Kornblith, H. (2001): *Epistemology: Internalism and Externalism*. (ed.) Oxford, Blackwell.

Lehrer, K. (1990): *Theory of Knowledge*: Boulder, Westview Press.

Pagano, M. and K. Gauvreau. (2000): *Principles of BioStatistics*. Duxbury, Australia.

Papineau, D (2003). *The Roots of Rationality*: Clarendon Press: Oxford.

Plantinga, A. (1993): *Warrant: The Current Debate*. Oxford, Oxford University Press.

Pollock, J. and J. Cruz. (1999): *Contemporary Theories of Knowledge*. Totowa, NJ: Rowman and Littlefield.

Psillos, S. (2004): "A Review of Achinstein's *The Book of Evidence*", *Philosophy and Phenomenological Research*. Pp. 740-743.

Quine, W. (1953): "Two Dogmas of Empiricism." in *From a Logical Point of View*. New York, Harper and Row.

Royall, R. (1997): *Statistical Evidence: A Likelihood Paradigm*. New York, Chapman Hall.

Royall, R. (2004): "The Likelihood Paradigm for Statistical Evidence". In Taper and Lele (2004) Chicago, University of Chicago Press, (eds.).

Salmon, W. (1983): "Confirmation and Evidence." In Achinstein (1983) (ed.)

Samuels, R., Stich, S., and Bishop, M. (2001), "Ending the Rationality War: How to make Disputes about Human Rationality Disappear", in R. Elio (ed.), *Common Sense, Reasoning and Rationality*, New York: Oxford University Press.

Skyrms, B. (1984): *Pragmaticism and Empiricism*. New Haven, Yale University Press.

Sosa, E. (1991): *Knowledge in Perspective*, New York, Cambridge University Press.

Swinburne, R. (2001): *Epistemic Justification*. Oxford, Clarendon Press.

Taper, M., and S. Lele. (2004): (eds.) *The Nature of Scientific Evidence*, Chicago, University of Chicago Press,.

Williams, M. (2001): *Problems of Knowledge*: Oxford, Oxford University Press.

Williamson, T. (2000): *Knowledge and its Limits*. Oxford, Oxford University Press.

Whewell, W. (1858): *The History of Scientific Ideas,* 2 vols. London, John W. Parker.

[1] See Heil (1992) on this point.
[2] Chisholm (1980). He seems to give the same impression when he provides "the theory of evidence from a philosophical—Socratic—point of view" in the chapter on "The Directly Evident" in his *Theory of Knowledge* (2nd edition) book.
[3] Pollock and Cruz (1999).
[4] Bernecker and Dretske (2000)
[5] Kahneman, Slovic, and Tversky (1982).
[6] Glymour (1980).
[7] All through this paper, we have assumed a crucial intuitive distinction generally overlooked in epistemological literature between "data" (or datum) and "evidence." This distinction is embraced by scientists working with large sets of data such as in statistics and data mining. For example, this distinction has an analogue in data mining in which data miners distinguish between "data" and "information." The idea behind the latter distinction is that we could be "data rich", but at the same time we could be "information poor" (see Han and Kamber 2001). For us, data is information extracted from nature and stored with some minimal retrieval structure. Evidence on the otherhand is data used within the context of a particular argument. Epistemologists need to make this distinction. Otherwise the notorious old evidence problem will be regarded as a trivial problem, for example, "does old evidence provide evidence for a new theory"? In fact, once we make the distinction between "data" and 'evidence", the old evidence problem poses a serious threat to theory testing, "do old data provide evidence for a new theory"?
[8] Achinstein (2001).
[9] See Royall(1997).
[10] Royall speaks only of belief, not degrees of belief.
[11] One might object that this might not be a representative probabilistic measure for Bayesian confirmation. One suggestion might be to use the measure Prob(H|D) > Prob(H) instead. To respond to this objection, we would like to point out first that there are also other measures, for example, the difference between Prob(H|D and Prob(H|~D), and many more. Second, we need to distinguish between two concepts of confirmation; the *absolute* confirmation and the *incremental* confirmation (see Earman and Salmon, 1992, p. 89, footnote 13 and Carnap, 1962). This notion of confirmation is intended to capture the absolute confirmation, while at the same time, making Bayesianism to be both minimally constrictive and maximally tolerant. In addition, Gill, a Bayesian sociologist, also endorses this absolute confirmation measure because, according to him, to judge the quality of a single model, there is no better measure than referring to their posterior summaries. (Gill, 2002, p. 271)
[12] Taper and Lele 2004 Chapter 16 define evidence as a data based estimate of the relative distance of two models from truth.
[13] Royall (1997)
[14] Lele (2004) Chapter 7 in Taper and Lele 2004
[15] Berger (1985) page 146. Often, the law of likelihood that provides justification for the use of the LR has been called into question. As a result, one might conclude from this that the BF measure must be called into question along with the LR. However, one should recall that the justification for the BF draws its strength from the likelihood principle and not from the LL. Since the likelihood principle is independent from the LL, so any attack on the latter does not necessarily constitute an attack on the likelihood principle.
[16] Royall (1997), Good (1983)
[17] In the comparison of complex hypotheses prior beliefs do not cancel as fully from the BF (see Berger 1985)
[18] Royall (1997) has provided an intuitive justification for the choice of number of 8 as a cut-off point for strong evidence. Royall has mentioned that this benchmark value (i.e., 8) or any value in that neighborhood is widely shared by statisticians of different stripes. Hence, it should come as no surprise that 8 or a number in that neighborhood is the benchmark. See Royall (2004) for more on this issue.
[19] Pagano and Gauvreau (2000).
[20] This point was drawn to our attention by Robert Boik. The proof of the theorem is as follows: BF > 1 ↔

Prob(H|D) > Prob(H) ↔ [Prob(D|H) Prob(H)]/ Prob(D)] > Prob(H) ↔ [Prob(D|H) > Prob(D)] ↔ Prob(D|H) > {[Prob(D|H) Prob(H)] +[Prob(D| ~ H) Prob(~H)]} ↔ 1 >[Prob(H) + Prob(~H)/BF] ↔ (1-Prob(H)) > (1-Prob(H))/BF ↔ BF > 1.

[21] Here we have adopted this formulation from Papineau (2003) with a minor change of his example.

[22] We are in the process of advancing an experiment that will test our hypothesis regarding the analysis of the base-rate fallacy.

[23] Perihelion is the time when a planet or comet is at its closet distance to the Sun. The orbits of these heavenly bodies are not usually circular, but rather they are elliptical and so there is a maximum distance (aphelion) and a minimum distance (perihelion). It has been observed that Mercury's position at perihelion (shifts a very small distance every rotation, A tiny portion of this shift, forty-three seconds of arc per hundred years, cannot be accounted for by classical mechanics including Newtonian theory. In contrast, GTR is able to account for this shift. So the perihelion shift of Mercury is taken as a test for GTR.

[24] Earman, J. 1992

[25] Glymour (1980),

[26] Glymour (1980), 86. Note, in this context T stands for theory, and Glymour is unclear about the distinction between evidence and data.

[27] Colin Howson has raised this objection in a personal communication with one of us.

[28] Jim Joyce in his APA, Eastern Divisional comments on one of our papers has raised this point against this BF/LR based account intended to solve the old evidence problem.

[29] Suppose for arguments sake that we could assign values to Prob(catch-all & M). There are three possible ways of explaining the dependence of Prob(GTR) on Prob(catch-all &M), and thus to defeat the objection that if the Prob(M) is high and the likelihood ratio is high (>1) then, Prob(GTR) has to be high. First, if the joint probability of the catch-all hypothesis i.e., Prob(catch-all & M), is reasonable, 0.4 or 0.5, then it does not make Prob(GTR) to be high. The other possibility could be that we just don't know the value of the above joint probability, because it involves the catch-all hypothesis. As a result, we can't say anything about Prob(GTR) being be high or low. The remaining possibility is in which the joint probability could be low, hence Prob(GTR) has to be high to satisfy (EQ4). Onely under this case, the above objection would arise. However, it does not follow, contrary to the objection, that in the old evidence problem, if the Prob(M) is high and the likelihood ratio is high, then the agent is bound to have high prior probability for GTR.

[30] Evidence is also independent of the intent of the investigator collecting the data, although it may depend on the way the data are collected We also do not intend to say that it is epistemologically irrelevant whether D was known before hypothesis formulation or not. These statements are discussed in greater detail in chapters 5, 7, and 16 of Taper and Lele (2004)*The Nature of Scientific Evidence*. University of Chicago Press. Chicago.

[31] Achinstein (2001), 24.

[32] Achinstein (2001),116.

[33] Achinstein (2001), 116

[34] The likelihood ratio here is the ratio of the binomial probabability of four heads out of five trials with a head's probability of 0.8 to the binomial probability of four heads out of five trial with a head's probability of 0.5 = Bin(4,5,0.8)/Bin(4,5,0.5)=0.4096/0.1562=2.62.

[35] Note, this evidence is the same regardless if one calculates the LR for the two models on the basis of 20 heads out of 25 throws, or as the product of the LRs for the five sequences.

[36] This example is adopted for our purpose from Pagono and Gauvreau (2000).

[37] See more on this in Dawid (2002).

[38] We borrow this title from Dennett (1987). However, his position is different from ours. Unlike Dennett, who would like to "salvage some theoretically interesting and useful versions of, or substitutes for, the concept of belief (117)", We would rather be inclined to add "evidence" as another central concept to epistemology along with belief.

[39] Royall (1997, 2004)

[40] Royall (1997, 2004)

[41] *The American Heritage Dictionary of the English Language* 4th edition (2000.

[42] The following example shows how this is possible. Consider an example where a general physician is confronted with a case in which a patient feels feverish and shows symptoms of malaria. However, no prior case of malaria has been reported in the locality or elsewhere. So if a person is chosen randomly, the probability that the person is afflicted with malaria must be very low. H is the hypothesis that the person has been afflicted with malaria. The doctor's prior probability for that person having malaria is assumed to be 0.09. Hence, Prob(~ H)=0.91. However, she has run some blood test on the patient to make sure whether he really has malaria. To her surprise, she has found out that the test is positive. She knows that the test is very reliable. Let D represents a positive screening test result and assume that Prob(D| H) = 0.95. Not all the people who are tested suffered from malaria. In fact, 1% of the tests are false positive outcomes. So Prob (D|~H) = 0.001. Now what's the probability that the individual in question has malaria given the test result is positive? To answer that, we have to compute Prob(H|D) = 0.98, which is close to 1. This is very high posterior probability. Hence, H must receive a strong confirmation. Should D be regarded as strong evidence for H as against its denial? The LR of Prob (D|H)/ Prob(D|~H), i.e., 11 times. This means that the probability an individual with a positive blood test has malaria is 11 times greater than the probability that an individual randomly selected from the population that has the disease. Here is a case in which the agent has both strong confirmation and strong evidence for the hypothesis that the individual is suffering from malaria.

One might contend that there is an inconsistency between our account of the resolution of the old evidence problem and above malaria example. This is how this alleged inconsistency argument goes. We said in response to the old evidence objection that to argue that if both the old data's probability is high and the BF is high, then the prior probability for the hypothesis for which the BF is high must be HIGH, is mistakenly asking a belief question, which is irrelevant for assessing the old evidence problem because the old evidence problem is raising an evidence question. We wrote that this charge (against our account) is comparable to the HIV case in which the first specialist is interested in the belief question, but if another specialist points out that the first specialist should be interested in an evidence question, then it would be asking an irrelevant question or providing an irrelevant recommendation. However, in the concluding part of the paper, we wrote that we never suggest that evidence could never provide reasons for believing a theory. We pointed out correctly that we suggested that they could very well overlap in some cases, e.g., the malaria example. The inconsistency argument is that that these two claims are inconsistent. Here is our response why they are not.

When somebody is asking a belief question, then one should be receiving an answer in terms of a confirmation account which is designed to address the belief question. At that time, if some one uses tools for an account of evidence to address that belief question, then the former is conflating the two questions. It is possible that when a confirmation account implies (let's assume) that one should believe the theory in question because Prob(H|D) > 0, an account of evidence could also imply that there is (strong) evidence for H as against ~H. The case in point is our TB example. However, when we are asking a belief question regarding what should we believe regarding whether the agent is suffering from TB, it is not (and it would be irrelevant if somebody expect it) necessary to know whether there is weak or strong evidence for TB as against its denial. The latter is a separate issue. There could be cases in which both strong confirmation for H and strong evidence for H could co-exist. The malaria example shows that how this is possible., In short, when one is asking a belief question, it is irrelevant to expect an account of evidence will provide that answer, although it is possible that both a confirmation account and an account of evidence could agree regarding what should be the recommendation at this point. For example, the former might say that one should believe the theory, when the latter might say that there is evidence (strong or weak, or very strong) for the hypothesis as opposed to its denial. So there is no inconsistency between our two claims, the first claim is made in connection with the old evidence problem and the second one is made in the concluding section of the paper.