ORIGINAL ARTICLE

Why Explanatoriness Is Evidentially Relevant

Kevin McCain[1] & Ted Poston[2]
[1] University of Alabama at Birmingham
[2] University of South Alabama

William Roche and Elliott Sober argue that explanatoriness is evidentially irrelevant. This conclusion is surprising since it conflicts with a plausible assumption — the fact that a hypothesis best explains a given set of data is evidence that the hypothesis is true. We argue that Roche and Sober's screening-off argument fails to account for a key aspect of evidential strength: the weight of a body of evidence. The weight of a body of evidence affects the resiliency of probabilities in the light of new evidence. Thus, Roche and Sober are mistaken. Explanatoriness is evidentially relevant.

We are explanationists. That is, we think that the fact that a hypothesis best explains some data is evidence that the hypothesis is true. This is a basic commitment of anyone who makes use of inference to the best explanation (IBE). Given the ubiquity of IBE in everyday life and the sciences, we find it surprising that William Roche and Elliott Sober (2013) have recently attempted to show that explanatoriness is evidentially irrelevant.[1]

Roche and Sober (2013, p. 659) claim that if explanatory considerations are evidentially relevant, then when considering some hypothesis H and some observation O, the proposition E, <If H and O were true, H would explain O>, should raise the probability of H. The notion of 'probability' here is epistemic probability, also referred to as 'rational credence'. In other words, they claim that explanatoriness is positively evidentially relevant if and only if $\Pr(H|O\&E) > \Pr(H|O)$. They argue, however, that this is *never* true because "O screens-off E from H" (2013, p. 660). That is, $\Pr(H|O\&E) = \Pr(H|O)$. On Roche and Sober's view, recognizing the fact that H explains O does not add anything to the degree to which O confirms H. We argue that the screening-off argument mistakes an evidential role explanatory considerations play. Explanatory considerations can make a Pr-function more resilient by making certain of its probabilities less volatile given future information. This volatility property, however, may not affect the probability of a hypothesis given certain evidence. We will thus grant that Roche and Sober specify a condition in which explanatory considerations are not evidentially relevant in their limited sense of evidential relevance.[2] However we show that there is a wider, more natural sense

Correspondence to: E-mail: mccain@uab.edu

of evidential relevance on which explanatory considerations are relevant to a subject's total evidence. Moreover, this wider notion of the strength of total evidence naturally fits within Bayesianism. Consequently, our thesis is that explanatoriness is evidentially relevant.

Before we turn to the screening-off argument, we note that Roche and Sober's argument is argument by a single case. That is, they argue that explanatoriness is *never* evidentially relevant because it is screened off when there is good observational data concerning the nonepistemic, objective chances relating smoking to cancer. We fail to see how this argument generalizes to *every* case. Roche and Sober rightly focus on a certain class of cases in which a subject has observational evidence regarding the nonepistemic, objective chance of a hypothesis and then press the objection that in that case explanatoriness is not evidentially relevant. Our reply to their screening-off argument is to show how explanatoriness is evidentially relevant in that condition. But there is a larger use of explanatoriness that their argument does not to impugn. For example, the ability of Newton's theory to explain the orbits of the planets is evidence that Newton's theory is true, even if we lack observational evidence regarding the nonepistemic, objective chance that Newton's theory is true. Similarly, the discovery that Einstein's theory of general relativity explained the precession of the perihelion of Mercury increased the probability of Einstein's theory. So, in these cases $\Pr(H|O\&E) > \Pr(H|O)$. One might fret about whether we can make sense of probabilities attending to general theories and then use those worries to mount an argument against the evidential relevance of explanatoriness, but Roche and Sober do not take that strategy. Even if one is not convinced by these remarks about how explanatoriness functions in other cases and inclined to think that explanatoriness is not evidentially relevant in these cases, read on because we show precisely how explanatoriness can increase the strength of evidence for a hypothesis in Roche and Sober's screening-off case.

1 The screening-off argument

Roche and Sober's screening-off argument appeals to frequency data concerning the correlation between smoking and cancer. Before scientists formulated any causal relation between smoking and cancer they observed that smokers tend to get cancer more frequently than nonsmokers and, further, that one's exposure to the risk of cancer increases the more cigarettes one smokes (2013, p. 660). It's natural to understand this frequency data as data concerning the objective, nonepistemic chance of a smoker developing cancer. According to Roche and Sober these frequency observations led scientists to justifiedly accept the following probabilistic inequality:

> (*) $\Pr(S$ will get lung cancer$|S$ has smoked i cigarettes to date$) > \Pr(S$ will get lung cancer$|S$ has smoked j cigarettes to date$)$, for all $i > j$. (2013, p. 660)

On a Bayesian analysis of confirmation this inequality implies that S being a heavier smoker than S^* is evidence that S has a greater chance of getting lung cancer than S^*. Different Bayesian measures of confirmation will assign some numerical value to the

degree of confirmation that rate of smoking makes to chances of cancer. For the purpose of our argument, let us say that the numerical value of the degree of confirmation this inequality supports is $\boldsymbol{d}$.[3]

Should adding an explanatory claim to one's background evidence given the same frequency data lead one to think that the degree of confirmation is any greater than $\boldsymbol{d}$? The frequency data was available prior to any formulation of a specific causal-cum-explanatory story about the relationship between smoking and cancer. Suppose we learn that there is a causal relationship between smoking and cancer such that we now fully accept that *smoking explains cancer*. That is, suppose we add to our background evidence the causal-cum-explanatory claim that greater exposure to smoking causally influences one's chance of getting cancer. Consider:

(**) Pr(S will get lung cancer|S has smoked i cigarettes to date & *smoking causes cancer*) > Pr(S will get lung cancer|S has smoked j cigarettes to date & *smoking causes cancer*), for all $i > j$.

We contend that *adding* the explanatory claim does not change the degree of confirmation between (*) and (**) and *yet* the explanatory claim increases the strength of total evidence. The degree of confirmation measured in (*) and (**) may both equal $\boldsymbol{d}$. That is, the degree of confirmation the rate of smoking makes to lung cancer can be *precisely* the same between having the frequency data alone and having that data *and* learning the explanatory claim. Even so, the explanatory claim significantly increases the total evidence the subject has for the inequality. How is this so? The explanatory claim *reinforces* the frequency data without *changing* the frequency data.

What is going on in the smoking and cancer case is that the frequency data provides good evidence about the relevant objective, nonepistemic chances relating smoking with cancer, evidence that thereby provides a good basis for assigning probabilities. When one learns the causal-cum-explanatory claim, one doesn't learn that the chances were wrong; rather one learns *why* the chances are as they are. In general, a theory about the relationship between two (or more) events need not change the observed frequencies of those events. Consequently, once the theory is added to one's evidence one shouldn't change one's beliefs about any specific case—e.g., the chance a heavy smoker will get cancer. Yet learning the theory does change the resiliency of those probabilities.[4] Whereas prior to possessing an explanation one only had observation to settle the frequencies, after gaining the explanation one has observation and theory. This is a standard case of a conclusion being over-determined and it should come as no surprise that, in such a case, learning an explanatory claim does not change the relevant frequencies.

The resiliency of a probability function reflects a dimension of strength of evidence that is not captured directly by first-order probabilities. Consider the following simple case to illustrate this dimension of evidential strength.[5] Suppose one has a two-sided coin, which looks fairly typical. What is the epistemic probability it lands heads on the next flip? If you knew the nonepistemic, objective chance of it landing heads on the next flip that should be the probability you assign. But failing that you must use your available evidence to determine the relevant nonepistemic, objective chance. Since your evidence indicates

that the coin is fair and the various bias hypotheses you entertain are symmetrically counterbalanced, you assign a value of 0.5 to the next flip resulting in heads. Now suppose one flips the coin a million times and it lands heads approximately one-half of the time. What is the chance that the next flip is heads? The answer is the same, 0.5. $\Pr_t$(heads on next flip$|k$) = 0.5 and $\Pr_{t+e}$(heads on next flip$|k$ & e) = 0.5, where k is one's background evidence, e is that there have been approximately $1/2$ heads among a million flips, $\Pr_t$ (--|--) is one's initial probability assignment and $\Pr_{t+e}$ (--|--) is the probability assignment after learning only e. Even though there's no change in the balance of probability, one's evidence for that assignment has increased significantly. How is this possible if the million flips *are screened off*? There's another dimension to a probability function: *resilience*. The resilience of a probability function concerns how it changes in response to new information. In the above case $\Pr_{t+e}$ is more resilient than $\Pr_t$. We contend that, under the conditions Roche and Sober provide, the explanatory claim increases the subject's strength of evidence even though it does not affect the relevant first-order probability. It makes that first-order probability more resilient. Roche and Sober write, "If you already know that O is true and you have computed $\Pr(H|O)$, learning E does not change how confident you should be in H" (2013, p. 660). This is only partially right. The first-order probabilities aren't affected; but the resilience of those probabilities is affected. Thus, there is a broader notion of strength of evidence according to which explanitoriness *is* relevant to a subject's total evidence.

Let us develop this response. James Joyce (2005) puts forward a more sophisticated Bayesian framework for understanding a subject's total evidence.[6] Joyce observes that Bayesians typically assume that the *total* amount of evidence for X is directly reflected in the probability of X (2005, p. 158). Joyce claims that Bayesians typically hold that (a) a person has more evidence for X than Y iff $\Pr(X) > \Pr(Y)$, for all admissible Pr-assignments, (b) a person has strong evidence for X iff $\Pr(X) \approx 1$, and (c) E provides a subject with incremental evidence for X iff $\Pr(X|E) > \Pr(X)$.[7] Joyce claims that "this picture of the relationship between credences and evidence is seriously misleading." We agree with Joyce. The correct picture needs to account for what Joyce calls the *weight* of evidence.[8] The weight of the evidence affects the resiliency of a Pr-function. Once this is done we see an evidential role for explanatoriness to play.

Joyce distinguishes between the balance, or direction of a subject's evidence, and the weight of that evidence. We will work with our original example to explain Joyce's distinction. Recall that in our case a subject has background evidence that indicates a given coin is fair and further that the evidence for various bias hypotheses is symmetrically counterbalanced. Thus, the subject assigns $\Pr(\text{Heads}|\text{flip}) = 0.5$. After flipping the coin a million times, the subject has much better evidence that the coin is fair, and thus, that the $\Pr(\text{Heads}|\text{flip}) = 0.5$. The *direction* or *balance* of the subject's evidence in both cases for the proposition that the next flip will be heads is the same. But the *weight* of the subject's evidence has significantly increased. The *new* evidence the subject acquires indicates that his *initial* evidence concerning the objective chance of the coin to land heads was correct. At the initial stage he could have easily acquired evidence that his credence in the objective chance of the coin comes apart from its actual objective chance. But after flipping

the coin a million times he has a much better evidential basis to judge its actual objective chance of landing heads. In both the initial stage and in the final stage the *direction* of his evidence is the same: i.e., the chance the coin will be heads is 0.5. But in the final stage his evidence decisively vindicates that the initial *direction* or balance of his evidence was correct.

The difference in the weight of the evidence occurs in the profile of the probability assignment updating on future information. Suppose in the initial stage our subject begins flipping the coin. He gets an improbable sequence of five heads in a row. At this point, the subject may revise his original estimate of the objective chance of the coin to yield heads, revising upwards. However, once the subject has a much more substantial weight of evidence, the estimate of the objective chance survives runs of misleading data. As Joyce says, "The weight of this evidence is reflected in the tendency for credences to stably concentrate on a small set of hypotheses about the proposition's objective chance" (2005, p. 176). On our view, this is precisely the way in which learning explanatory information can play a role in the condition of screening-off. Explanatory information fixes the data about the relevant objective chances. It does not change what the data says the relevant objective chances are. Thus, if one has frequency data about the relevant objective chances, that is a feature of one's evidence that indicates what the objective chance is (obviously). But, once one acquires an explanatory-cum-casual story that indicates that the objective chances are the same as the frequency data shows, the *weight* of the evidence is significantly increased even though in both cases the *direction* of the evidence is the same. Adding the explanatory story significantly changes the probabilistic profile of the Pr-function when updating on future information.

It may be helpful at this point to consider a toy example to see how explanatory considerations can render a probability function more resilient. Consider the following case: Sally and Tom have been informed that there are 1,000 x-spheres in an opaque urn. Sally and Tom have the same background evidence except for this difference: Sally knows that blue and red x-spheres must be stored in exactly equal numbers because the atomic structure of x-spheres is such that if there are more (or less) blue x-spheres than red, the atoms of all of the x-spheres will spontaneously decay resulting in an enormous explosion. Sally and Tom observe a random drawing of 10 x-spheres without replacement, 5 blue and 5 red. The x-spheres are replaced in the urn. Given the data both Sally and Tom should assign $Pr(\text{blue}|\text{random draw}) = 0.5$. Sally has a very good explanation for why the probability of drawing a blue x-sphere at random is .5, but Tom only has the frequency data to go on. Yet this evidential difference does not show up given the initial data. Suppose, however, that 10 more x-spheres are drawn at random and they are all blue. Sally's rational credence in blue given a random draw remains the same. $Pr_{\text{Sally}+E}(\text{blue}|\text{random draw}) = .5$, where this is her revised probability given learning just the result of the new draw, E. Yet Tom's rational credence changes significantly. Whereas $Pr_{\text{Tom}}(\text{blue}|\text{random draw}) = .5$ for the initial round of testing, $Pr_{\text{Tom}+E}(\text{blue}|\text{random draw}) = .75$ when he learns just E. So, Tom's probability function is more volatile in the face of the new information. Sally's probability function is more resilient.

Our primary response to Roche and Sober's screening-off argument is complete. Explanatoriness is evidentially relevant even in the screening-off condition by increasing the total weight of evidence. In the next section we offer some initial replies to additional arguments Roche and Sober offer to undermine the claim that explanatoriness is evidentially relevant.

2 Replies to additional concerns

In the following we consider three additional concerns Roche and Sober raise about the evidential role of explanatoriness.

2.1 Is observation enough?

Roche and Sober maintain that observation is enough to settle a rational credence in a hypothesis. If correct, this would show that when we have observational data explanatory considerations are superfluous. Fortunately, Roche and Sober do not successfully establish this claim because their example requires a causal-cum-explanatory background. The frequency data that Roche and Sober allude to is data scientists assembled concerning the relationship between smoking and lung cancer. This assemblage of data gradually produced a basis for justified beliefs about the objective, nonepistemic probabilistic relationship between smoking and lung cancer. Why is it that the data had this feature that ultimately allowed scientists to have justified beliefs about the relevant correlation? The data was not widely divergent; rather the data indicated that there was some causal process — albeit unknown at that time — that explains the correlation between smoking and lung cancer. If it turned out that there was *no* explanatory story to be told about the data then that would be *extremely surprising*. Exactly this feature — a justified belief in an unknown explanatory story — plays a crucial role in using the data from observation to get justified beliefs about the relevant frequencies. Apart from a general justified belief in some explanatory story accounting for that data, the observational data would not justify beliefs about the relevant frequencies. In this case the data would not be inductively projectable. The key feature of Roche and Sober's view is that the data is projectable and that provides the basis for the screening-off claim. But on our view one cannot get inductive projection to unexamined cases apart from explanatory considerations.[9] As we understand it, this is the lesson of the grue problem. Grue-like regularities are not projectable because they are not explanatory. On our view, some of the plausibility of Roche and Sober's view depends on assuming a broader explanatory context for the observational data. Apart from an explanatory framework, the observational data would not be projectable; thus, the observational data wouldn't even support the inequality established by the frequencies of incidents of smoking and cancer. Thus, even if one grants Roche and Sober's claim that "O screens-off E from H" this doesn't show that explanatory considerations are irrelevant to confirmation. Explanatory considerations are already at work in setting $\Pr(H|O)$ — having E provide additional confirmation for H would be akin to double-counting the information about objective chances. While double-counting is a mistake, giving explanatory considerations their proper due is not.

2.2 Explanation and confirmational symmetries

An additional concern for the evidential relevance of explanatory considerations arises from the brief just-so story that Roche and Sober offer to account for the failure of explanatoriness to be evidentially relevant. They observe that explanation is antisymmetric but confirmation is symmetric, and so claim that their thesis is not at all surprising given this difference. They write,

> If the explanation relation is anti-symmetric and the evidence relation is symmetric, it is no surprise that evidential relations are sometimes indifferent to explanatory relations. If smoking is evidence for lung cancer, then lung cancer is evidence for smoking, and it makes no difference that smoking explains lung cancer but lung cancer does not explain smoking. (2013, p. 662)

They claim that it is not implausible that the frequency data supports the following:

> (+) $\Pr(S$ smoked cigarettes earlier in life$|S$ gets lung cancer later$) = \Pr(S$ gets lung cancer later in life$|S$ smoked cigarettes earlier$)$. (p. 662)

However, they maintain that if explanatoriness is evidentially relevant, then the following inequality should be true:

> (++) $\Pr(S$ smoked cigarettes earlier in life$|S$ gets lung cancer later & earlier smoking would explain later lung cancer$) > \Pr(S$ gets lung cancer later in life$|S$ smoked cigarettes earlier & later lung cancer would not explain earlier smoking$)$. (p. 662)

But, Roche and Sober claim that this inequality is not supported by the frequency data. So, they conclude that explanatoriness is not evidentially relevant. However, the inference from the lack of support for (++) to the claim that explanatoriness is not evidentially relevant is not warranted. Let us explain.

First of all, explanationists are not committed to the antisymmetry of explanatory relations leading to antisymmetry in confirmation. As Gilbert Harman (1973) argues, explanationists can plausibly maintain that *Y confirms X* when *X* explains or is explained by *Y*. After all, explanationists hold that confirmation comes from being part of the best overall explanatory picture—propositions can be part of this picture by explaining or being explained. So, contra Roche and Sober, it would be surprising if their thesis were true. This, of course, does not show their thesis is false, but it is worth noting because it makes it clear that it is not simply the sort of thing that we should expect to be true given what we already know about explanatory relations and confirmation.

Second, for reasons we gave above relating to resilience, explanationists have a principled reason for denying (++). On our view (++) is false because (+) is true and the fact that *smoking explains cancer* merely makes those probabilities more resilient. Thus, the fact that *later lung cancer does not explain earlier smoking* is irrelevant. Where we have justified frequencies we should go with those; explanatory considerations are relevant because they make those frequencies more robust.

2.3 First priors

A final concern for our position comes from consideration of first priors. We have argued that there is an evidential role for explanatory considerations to play at the level of resilience and that moreover explanatory considerations are required for inductive projection. The necessity of explanatory considerations for inductive projection dovetails into a view of the role of explanatory considerations defended by several authors.[10] Roche and Sober claim that their reasoning that "closes the door for degree of confirmation also slams it shut for prior probabilities" (2013, p. 666). They reason that since today's priors are often yesterday's posterior probabilities, the reasoning they give shows that explanatory considerations are evidentially irrelevant. When one has justified frequency beliefs, explanatoriness is screened off. It gets screened off for yesterday's posterior probability and so explanatoriness doesn't affect today's prior probability. We have, however, explained the mistake in this argument. Explanatoriness is evidentially relevant even when screened off because it makes probabilities more resilient.

Roche and Sober then consider whether explanatory considerations should be reflected in the choice of initial prior probabilities. They ask: "Should first priors be assigned values to reflect considerations of explanatoriness? We are sceptical" (2013, p. 666) Roche and Sober then proceed with some parting shots to indicate the direction of their skepticism. For instance, they write, "First priors are supposed to be assigned on the basis of zero observational information. We are okay with tautologies and contradictions being assigned priors in this circumstance though this has nothing to do with explanatoriness" (2013, p. 666). We agree that there is a research project here and there are questions to be answered, but we are optimistic that the evidential role of explanatoriness can be vindicated. Huemer (2009) argues that explanatory considerations are required for inductive projection. If one does not appeal to the explanatoriness of objective chances, it seems that one should distribute prior probabilities over sequences of events in such a way that inductive scepticism follows. As Huemer argues, explanatory considerations are required for a prior probability distribution that allows for inductive confirmation to unexamined instances.[11] Weisberg (2009) argues that explanatory considerations can be used to make the principle of indifference more tractable. First priors should be distributed over hypotheses in a way that reflects which hypotheses are more fundamentally explanatory than others. For instance, if one takes the side length of a cube to be explanatorily prior to its volume then the side length of a cube is the parameter one should use for some suitable indifference principle. Using explanatory depth to fix relevant parameters for the principle of indifference answers some of the inconsistency objections to the indifference principle. Once fundamental explanatory parameters are fixed one should distribute confidence in a way that reflects the simplicity of the relevant hypotheses. Of course, there are challenging details to work out on this approach but this does not show that, in general, explanatoriness isn't evidentially relevant. More to the point, though, the general challenges that Roche and Sober point to are distinct from their frequency-based argument against explanatoriness. We have argued that that argument is undermined by the role explanatoriness can play with respect to resilience. Moreover, we've argued that

some explanatory features are required for observed frequencies to be used for inductive projection.

Acknowledgment

Notes

 1 See Glymour (1984) and Thagard (1978) for discussion of the prevalence of IBE in our everyday reasoning and examples of it being employed in the sciences.
 2 We leave open the issue of whether explanatoriness is screened off in every case in which Roche and Sober's condition is not met.
 3 The specific choice of a measure of confirmation does not affect the argument in the text. See Fitelson (1999) for a discussion of measures of confirmation.
 4 See Skyrms (1977) for a discussion of resiliency.
 5 Discussed by Skyrms (1977, pp. 705–6).
 6 Thanks to Branden Fitelson for bringing to our attention the relevance of Joyce's paper to our claims.
 7 Joyce (2005, p. 158). We have made slight changes to Joyce's notation.
 8 Joyce sources his remarks about weight in Skyrms as well.
 9 See Huemer (2009) and Poston (forthcoming) for a defense of this claim.
10 Huemer (2009), Weisberg (2009), and Poston (forthcoming).
11 See also Poston (forthcoming) for a defense of this claim.

References

Fitelson, B. "The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity." *Philosophy of Science* **66** (1999): S362–78.

Glymour, C. "Explanation and Realism," in *Scientific Realism*, edited by J. Leplin. Berkeley: University of California Press, 1984, 173–92.

Harman, G. *Thought*. Princeton: Princeton University Press, 1973.

Huemer, M. "Explanationist Aid for the Theory of Inductive Logic." *British Journal of Philosophy of Science.* **60** (2009): 345–75.

Joyce, J. M. "How Probabilities Reflect Evidence." *Philosophical Perspectives* **19** (2005): 153–78.

Poston, T. *Reason & Explanation: A Defense of Explanatory Coherentism*. New York: Palgrave-MacMillan, Forthcoming.

Roche, W. and E. Sober. "Explanatoriness Is Evidentially Irrelevant, or Inference to the Best Explanation Meets Bayesian Confirmation Theory." *Analysis* **73** (2013): 659–68.

Skyrms, B. "Resiliency, Probability, and Causal Necessity." *Journal of Philosophy* **74** (1977): 704–13.

Thagard, P. "The Best Explanation: Criteria for Theory Choice." *Journal of Philosophy* **75** (1978): 76–92.

Weisberg, J. "Locating IBE in the Bayesian Framework." *Synthese* **167** (2009): 125–43.