

The Evidential Impact of Explanatory Considerations

Kevin McCain & Ted Poston

Explanationism is an attractive family of theories of epistemic justification.¹ In broad terms, explanationism is the idea that what a person is justified in believing depends on their explanatory position. At its core, explanationists hold that the fact p would explain a body of evidence if p were true is itself evidence that p is true. In slogan form: explanatoriness is evidentially relevant. Although explanationism has been out of the limelight for quite a while, there has been a resurgence of interest in these views. We hope that this resurgence is in part due to our recent work defending explanationist views, but even if not, we will be happy if such views continue to garner the attention that they deserve.²

Despite the plausibility of explanationism, not all of the recent interest in it has been complimentary. Recently, William Roche and Elliott Sober (2013 & 2014) have argued that “*explanatoriness is evidentially irrelevant*” (2013: 659).³ R&S’s argument against the evidential relevance of explanatory considerations begins with what they call the “Screening-Off Thesis” (SOT):

Let H be some hypothesis, O be some observation, and E be the proposition that H would explain O if H and O were true. Then O screens-off E from H: $\Pr(H | O \& E) = \Pr(H | O)$. (2014: 193)⁴

R&S contend that *SOT is true if and only if “explanatoriness is evidentially irrelevant”*. We refer to this claim as (IRRELEVANCE). Putting these two together yields R&S’s conclusion that “*explanatoriness is evidentially irrelevant*”.

R&S’s conclusion is surprising and far-reaching. As we noted above, explanationism is an attractive view of justification. After all, it seems clear that a proposition’s being part of the best explanation of a body of data is itself a reason to think that the proposition is true. However, R&S’s conclusion, if correct, would spell serious trouble for explanationism. What is more, R&S’s argument does not simply threaten explanationism as a theory of justification, it threatens the viability of inference to the best explanation. If explanatoriness is evidentially irrelevant then we cannot be justified in inferring the truth of a proposition on the basis of its being part of the best explanation of a set of facts. This implies that Charles Darwin in *The Origin of the Species* actually gives a bad argument when he defends the use of explanatory reasoning in science:

¹ Hereafter, we will drop the qualifier “epistemic”, but any use of “justification” or its cognates should be understood to mean epistemic justification unless specified.

² See, for example, McCain (2014), (2016) and Poston (2014).

³ Henceforth, we refer to Roche and Sober as ‘R&S’.

⁴ The relevant notion of probability in SOT and throughout this discussion is epistemic probability. This sort of probability is often referred to as “rational credence”. It is because of this that we will often transition back and forth from speaking of probabilities to speaking of credences. Our hope is that where we do so the discussion is made clearer by our using either the term “probability” or “credence”.

It can hardly be supposed that a false theory would explain, in so satisfactory a manner as does the theory of natural selection, the several large classes of facts above specified. It has recently been objected that this is an unsafe method of arguing; but it is a method used in judging of the common events of life, and has often been used by the greatest natural philosophers. (1872/1962: 476)

Similarly, if IRRELEVANCE is true, Antoine Lavoisier provides poor support for his oxygen theory of calcification and combustion when he appeals to its explanatory virtues:

I have deduced all the explanations from a simple principle, that pure or vital air is composed of a principle particular to it, which forms its base, and which I have named *the oxygen principle*, combined with the matter of fire and heat. Once this principle was admitted, the main difficulties of chemistry appeared to dissipate and vanish, and all the phenomena were explained with an astonishing simplicity.⁵

If R&S are correct then the explanatory considerations that Darwin and Lavoisier track in their momentous scientific work provide no evidence that their theories are true. Furthermore, the widespread and common use of inference to the best explanation from motorcycle repair to administrative planning to medical diagnosis is misguided.

Fortunately, R&S are incorrect, and their argument fails. In a recent article (2014), we argued that R&S overlook an important dimension of evidential support when making their case that IRRELEVANCE is true, viz., the resilience of a probability function. Resilience is essentially how volatile a probability function is with respect to new evidence; a probability function with low volatility is more resilient than a function with high volatility. We maintained that IRRELEVANCE is false because there are clear cases where explanatory considerations increase the resilience of a probability function. Additionally, we argued that there are numerous cases where SOT fails to hold.

R&S (2014) argue that we were mistaken on both accounts. The arguments R&S offer are not persuasive, but they do significantly clarify the disagreement. We pick up on this improved dialectical situation to further defend our position that explanatoriness is evidentially relevant. The upshot of our discussion is that both SOT and IRRELEVANCE are false because explanatory considerations may be captured in logical and mathematical relations encoded in a Pr-function. Thus, both inference to the best explanation and explanationism are safe from attack.

1. *Explanatoriness matters even if SOT is granted*

We provided a case to illustrate that IRRELEVANCE is false (2014: 149). The case shows how explanatoriness can make an evidential difference by making a probability function more resilient even if SOT is granted. In this case we focused on two subjects, Sally and Tom. Both are aware that there are 1,000 “x-spheres” in an opaque urn. Sally and Tom both observe the same random drawing (with replacement) of 10 x-spheres. An equal number of blue x-spheres and red

⁵ (Paris: Imprimerie Imperiale, 1862) Vol II, p. 623 (Thagard’s translation)—quotes here from Thagard (1978)).

x-spheres are drawn. Sally and Tom have the same updated credences after observing the outcome of the experiment. Namely, they both assign a probability of 0.5 to the next x-sphere randomly drawn being blue. However, there is an important difference between Sally and Tom. Sally, but not Tom, has an explanation of why the experiment turned out the way it did. As we explained, “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” (2014: 149). We pointed out that the explanatory difference between Sally and Tom makes Sally’s credence more resilient to future misleading information. That is, Sally’s credence remains 0.5 that the next x-sphere randomly drawn will be blue even given an improbable run of 10 blue x-spheres, while Tom’s credence significantly changes in response to this new (misleading) information. Thus, we concluded that even if it is granted that SOT is true, and so Sally’s having an explanation is screened-off from raising her credences in various outcomes, the explanatory difference between Sally and Tom makes an evidential difference. Thus, IRRELEVANCE is false.

R&S (2014) agree that Sally’s credence is more resilient than Tom’s in our case, but they deny that the reason for this difference is that Sally has an explanation and Tom does not. They highlight four features of our case:

- (i) Sally and Tom have a credence of 0.5 in proposition H (that the x-sphere drawn on the next random draw will be blue), (ii) Sally’s credence is more resilient than Tom’s, (iii) Sally but not Tom knows that if blue and red x-spheres are stored in unequal numbers, then there will be an enormous explosion, and (iv) Sally but not Tom has an explanation of why the probability of the blue x-sphere on a random drawing from the urn is 0.5. (p. 196)

R&S write, “It is true that Sally but not Tom has an explanation for why the probability of a blue x-sphere on a random drawing from the urn is 0.5, but this difference between Sally and Tom is doing no work” (p. 196). Rather, on their view, what does all the work is (iii). They explain, “Our point is that (ii) is true because (iii) is true, and (iv) does nothing to make (ii) true once (iii) is taken into account” (p. 197). This is to say that (iv) is screened off once (iii) is taken into account. On their view, explanatoriness is evidentially irrelevant in this case because Sally’s knowledge concerning the atomic structure of x-spheres makes (ii) true.

This move fails. In our example, the property of *Sally having the explanation* is one and the same property as the property of *her knowing the relevant fact about x-spheres*. In order to see this, consider an argument parallel to R&S’s that denies that water has the property of *extinguishing fire*. Water is H₂O, and H₂O has special chemical properties that make it an excellent chemical to extinguish fire. Because of the high degree of hydrogen bonding between water molecules, H₂O has the second highest heat capacity of all known substances. In virtue of its high heat capacity, its transition from a liquid to a gas requires a significant amount of energy, which enables it to rapidly quench flames. Once we account for these facts about H₂O the fact that water is present is screened off. Thus, water is irrelevant to extinguishing fire because H₂O is doing all the work.

The response to the water/H₂O argument is obvious: Water = H₂O. The properties of H₂O in virtue of which it makes an excellent fire extinguisher are the properties of water. They are one and the same. The same consideration applies to our x-sphere case. We might put our point thusly: to the extent that water has the property of extinguishing fires, explanatoriness in

our x-sphere case is evidentially relevant. Water extinguishes fire in virtue of its chemical structure. Explanatoriness is evidentially relevant in virtue of it specifying certain relations between H and E that get encoded in a Pr function.

It may be replied that the parallel argument breaks down because water is identical to H₂O, but Sally's having the explanation is not identical to her knowing the relevant facts about x-spheres. R&S may think that Sally can know that blue and red x-spheres are stored in equal numbers to prevent an enormous explosion without having an explanation of why the probability of a blue x-sphere on a random drawing from the urn is 0.5. This claim, however, is contentious, and R&S do not provide support for it. The claim is clearly false given a causal account of explanation. If knowledge of causal relations constitutes knowledge of explanatory relations, as causal accounts of explanation hold, then knowledge of causes is necessary and sufficient for explanatory knowledge. In which case, (iii) and (iv) are describing the same state of affairs.

Our initial (2014) reply to R&S left open the nature of the explanatory relation, but as causal accounts of explanation make clear, if explanations are constituted by such causal facts then conditionalization on causal facts will screen-off explanatoriness. But this is just the sense in which water is screened off from H₂O. In other words, it is not really screened off at all because those facts have already been taken into account. Conditionalizing separately on Sally's knowledge of the causal relation (as expressed in (iii)) and her knowledge of the explanation (as expressed in (iv)) is counting the same facts twice.

R&S do acknowledge that explanatoriness may be evidentially relevant if it is indicative of some fact, *I*, which specifies a probabilistic relation between H & E. *I* might be the proposition that H entails E, for example. In such a case R&S allow that if explanatory considerations in some way indicates to us that *I* is true, then such considerations may be evidentially relevant. However, they point out that Bayesian confirmation theory assumes agents are logically omniscient, and such omniscient agents would know logical and probabilistic facts like *I*. Thus, they maintain "explanatoriness has no confirmational significance, once purely logical and mathematical facts are taken into account" (p. 195).

We have two replies to this move. First, this response assumes that explanatory considerations are entirely separate from probabilistic relations between a hypothesis and evidence. If knowing the explanatory facts amounts to knowing some probabilistic relations between a hypothesis and evidence, then this response fails. It is exactly this sort of thing that we argue occurs in our x-sphere case—knowledge of the causal facts *just is* knowledge of the explanatory facts. In terms of the previous paragraph this means that in some cases knowledge of *I* just is knowledge of the explanatory considerations. So, there is no way to take *I* into account without taking explanatory considerations into account. Thus, we have good reason to think that IRRELEVANCE is false.

Second, there is a larger issue about what counts as 'logical and mathematical' facts. Is it a logical or mathematical fact that water extinguishes fire? If a person had complete knowledge of chemical theory could they deduce that water extinguishes fire in a normal environment (or, that the probability of its doing so was high)? Arguably, yes. But if these kinds of facts are facts a logically omniscient agent is aware of, then clearly once they are taken into account explanatoriness is evidentially irrelevant. Again, this is because the logical/mathematical facts encompass the explanatory facts. So, explanatoriness appears evidentially irrelevant, but this is simply because those facts have already been counted. Thus, if this is how one should

understand R&S's SOT thesis, then it amounts to the prohibition of counting the same facts twice. This should not concern any explanationist.

2. *Why SOT does not hold in all cases*

In addition to arguing that explanatoriness is evidentially relevant even in cases where it is granted that SOT holds, we argued that SOT is false in a number of cases. Here is the particular cases we mentioned in our earlier response to R&S:

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 $\Pr(H | O \& E) > \Pr(H | O)$ (p. 146).

R&S maintain that we are mistaken on this point. They claim:

Suppose you know that a theory H (e.g. Newton's) logically implies O (given the background information codified in Pr) and so you realize that $\Pr(O | H) = 1$. You then work out the value of $\Pr(H | O)$ by first obtaining values for $\Pr(H)$ and $\Pr(O)$. It follows (if neither H nor O has a probability of 1 or 0) that $\Pr(H | O)$ is greater than $\Pr(H)$. You then learn that O and as a result increase your credence in H. Suppose you later learn E. M&P's view [that is, our view] entails that upon learning E you should *further* increase your credence in H. It seems clear, however, that you should not do this. (p. 195)

R&S maintain that even in cases like those depicted by us explanatory considerations are not playing an evidential role. Rather, they maintain that all of the evidential work is being done by logical and mathematical truths. Thus, R&S maintain that we failed to show that there are cases where SOT does not hold.

One thing that explanationists sympathetic to our position could immediately question is how the values for $\Pr(H)$ and $\Pr(O)$ are obtained. The explanationist could plausibly argue that setting such prior probabilities will require an appeal to explanatory considerations. Plausibly, if these priors are set without appealing to explanatory considerations, the door to inductive skepticism will be left wide open. Huemer (2009), Poston (2014), and McCain (2016) each argue that the threat of inductive skepticism can only be avoided by way of explanatory reasoning. That is to say, without appealing to explanatory considerations we cannot establish a probability distribution that will allow for inductive confirmation. We mention this because it is a reasonable point for explanationists to press against R&S, but we will not pursue this line of argument here. Instead, we will point out the more glaring problem with R&S's argument.

The mistake that R&S are making here is similar to the mistake we pointed out in the first section. R&S are assuming that a particular kind of knowledge is distinct from having an explanation—they are assuming that coming to know that H entails O is not explanatory knowledge. This assumption, at least in the cases they are discussing, is mistaken. It is exceedingly plausible that knowing that H entails O (especially when H is a purported natural

law such as we have in Newton's theory) is the same as having an explanation of O , assuming that H is true. After all, this is part of the Deductive-Nomological (D-N) model of explanation.⁶ The D-N model does not work in all cases (e.g., flagpoles & shadows, hexed salt, and so on), but it does work in the sort of case that R&S describe where a natural law predicts/entails an observation. This is arguably a paradigm case of the sort that Hempel and Oppenheim had in mind when formulating this theory of explanation.

Apparently, R&S assume that an explanation must in all cases be something beyond knowing that H entails O . As they say in their original (2013) discussion of this topic, "even if proposition (E) has entailments about the logical and probabilistic relations of O and H , there is more to explanation than this...reasonable candidates for that something more will vindicate our screening-off thesis" (p. 663). It seems that R&S see their screening-off argument as similar to van Fraassen's (1989) Dutch book argument against inference to the best explanation. Both arguments rely on the assumption that in order to be evidentially relevant explanatory considerations must be something over and above logical and probabilistic relations. But, this is a mistake. While there are clearly cases where more is desired by an explanation, there are contexts where information about logical and probabilistic relations is explanatory. For instance, explanations in pure mathematics will plausibly consist entirely of information concerning logical and probabilistic relations. Insofar as explanations in pure mathematics are genuine explanations information about logical and probabilistic relations will be sufficient for explanation in at least some cases.⁷

It is worth briefly summing up a few points here. We agree with R&S that it would be a mistake to increase your credence in Newton's theory (H) upon learning E (H explains some observation, O) in the case as they describe it. In their discussion of the case R&S assume that you have already increased your credence in H on account of learning O . They assume that you possess the correct value for $\Pr(O|H)$ and $\Pr(H)$; so by Bayesian learning you update $\Pr(H)$ by $\Pr(H|O)$. You then learn E , which amounts to realizing that the probabilistic relation encoded in the value to $\Pr(O|H)$ is an explanatory relation. In this case it is clear that you should not further increase your credence upon learning E because you have already taken account of E when you include your knowledge of $\Pr(O|H)$. The evidential relevance of explanatory considerations does not license double counting. Thus, it seems that the case of Newton's theory is plausibly a case in which SOT does not apply. Similar considerations apply to the case we described concerning Einstein's theory of relativity as well as purely mathematical explanations in general.

Perhaps R&S think that in the cases of Newton's theory, Relativity, and pure mathematical explanations one does not actually possess an explanation unless she has a particular psychological state of understanding, an "aha!" moment. It may be that they think that this psychological state is required for explanation, and yet at the same time when a person has the "aha!" moment it adds nothing to confirmation.

While R&S may be correct that such a state does not further increase the probability of the relevant hypothesis (though such an experience may increase resiliency and so be evidentially relevant), it is a mistake to think that such a psychological state is required for possessing an explanation. It is one thing to have the information, and hence the evidential confirmation

⁶ Hempel and Oppenheim (1948)

⁷ For more on mathematical explanations see Hafner and Mancosu (2005), Kitcher (1989), and Steiner (1978).

provided by an explanation, and it is another thing to have this particular psychological state.⁸ So, attempting to argue that information about logical and probabilistic relations fails to be explanatory because it may not include an “aha!” moment is not promising. There are contexts where such information is explanatory whether or not it is accompanied by a particular feeling of understanding.

R&S mistakenly think that knowing that E (if H and O were true, H would explain O) is something beyond knowing that H and O are true and that H entails O because explanation is not a purely logical fact (2013: 664). This assumption coupled with some basic considerations from Bayesian confirmation theory lead R&S to wrongly claim that explanation is evidentially irrelevant. Bayesian confirmation theory assumes that agents are logically and mathematically omniscient. R&S play along with this but note that it can give the wrong results in some cases. They give the following case. Consider some proposition *I*, which says that O logically implies H. *I* can be evidence for H. Suppose you know that O is true, but you don’t know *I*. You then learn *I*. Your credence in H should change (assuming you were not already certain that H is true). Yet on the assumption of logical omniscience, *I* is not evidence. *I* is screened off from H since $\Pr(H|O) = \Pr(H|O \& I)$. (See R&S (2013), p. 664.) R&S hold that in this case the screening-off test “isn’t a good test for the confirmation relevance of purely logical facts” (2013: 664). But they think it is fair game for explanatory facts because explanatory facts are not logical facts.

The crucial question then is whether explanatory facts are, at least at times, indicative of facts about full or partial entailment (probabilistic facts are often thought of as facts about partial entailment). If E is so indicative, then the screening-off test is not a good test for confirmational relevance. In our case of Sally and Tom described above E is indicative of a probabilistic fact because the explanation specifies that the probability of random selection must be 0.5. Thus, the screening-off test is not a good test for the confirmational relevance of E in our x-sphere case. Moreover, often explanations work by specifying some previously unknown fact—e.g. the fact that space-time is curved. R&S (2014) discuss this in some detail. They acknowledge that there can be cases where E increases the probability of some hypothesis conditional on O: $\Pr(H|O \& E) > \Pr(H|O)$, especially where E is indicative of a logical fact like *I*. But they claim “explanatoriness has no confirmational significance, once purely logical and mathematical facts are taken into account” (2014, p. 195). This brings us back to our earlier criticism. To the extent that water has the property of extinguishing fire, explanatoriness has the property of being evidentially relevant. On our view R&S’s screening-off test amounts to a prohibition against counting the same facts twice. While good methodological advice, it does nothing to show that explanatoriness is evidentially irrelevant.

In light of the considerations adduced here it is reasonable to conclude that both SOT and IRRELEVANCE are false. Explanatoriness is evidentially relevant.⁹

References

⁸ For more on the relation between explanation and the psychological sense of understanding see de Regt (2004), Grimm (2009), and Trout (2002), (2005).

⁹ We are grateful to Kenny Boyce, Nevin Climenhaga, Brad Monton, Mark Newman, and Bill Roche for helpful comments on and discussion of earlier drafts.

Darwin, C. *The Origin of Species. 6th Edition.* New York: Collier, 1962.

de Regt, H.W. “Discussion Note: Making Sense of Understanding.” *Philosophy of Science* 71 (2004): 98-109.

Grimm, S. “Reliability and the Sense of Understanding”. In H.W. de Regt, S. Leonelli, & K. Eigner (eds.), *Scientific Understanding: Philosophical Perspectives*, Pittsburgh, PA: University of Pittsburgh Press (2009): 83-99.

Hafner, J. and P. Mancosu. “The Varieties of Mathematical Explanation.” In P. Mancosu *et al.* (eds.), *Visualization, Explanation and Reasoning Styles in Mathematics*, Berlin: Springer, (2005): 215–50.

Hempel, C.G. and P. Oppenheim. “Studies in the Logic of Explanation.” *Philosophy of Science* 15 (1948): 135-75.

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

Kitcher, P. “Explanatory Unification and the Causal Structure of the World.” In P. Kitcher and W. Salmon (eds.), *Scientific Explanation*, (Minnesota Studies in the Philosophy of Science, Volume XIII), Minneapolis: University of Minnesota Press, (1989): 410–505.

McCain, K. *Evidentialism and Epistemic Justification.* New York: Routledge, 2014.

McCain, K. *The Nature of Scientific Knowledge: An Explanatory Approach.* Switzerland: Springer, 2016.

McCain, K. and T. Poston. “Why Explanatoriness is Evidentially Relevant.” *Thought* 3 (2014): 145-53.

Poston, T. *Reason & Explanation: A Defense of Explanatory Coherentism.* New York: Palgrave-MacMillan, 2014.

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

Roche, W. and E. Sober. “Explanatoriness and Evidence: A Reply to McCain and Poston.” *Thought* 3 (2014): 193-99.

Thagard, P. “The best explanation: Criteria for theory choice” *Journal of Philosophy*, 75 (1978): 76-92.

Trout, J.D. “Scientific Explanation and the Sense of Understanding.” *Philosophy of Science* 69 (2002): 212-33.

Trout, J.D. "Paying the Price for a Theory of Explanation: de Regt's Discussion of Trout." *Philosophy of Science* 71 (2005): 198-208.

Steiner, M. "Mathematical Explanation." *Philosophical Studies* 34 (1978): 135-51.

van Fraassen, B. *Laws and Symmetry*. Oxford: Oxford University Press, 1989.