DOI: 10.1111/nous.12469

ARTICLE

NOÛS

A false dichotomy in denying explanatoriness any role in confirmation

Marc Lange

The University of North Carolina at Chapel Hill

Correspondence
Marc Lange, The University of North Carolina at Chapel Hill.
Email: mlange@email.unc.edu

Abstract

Roche and Sober (2013; 2014; 2017; 2019) have offered an important new argument that explanatoriness lacks confirmatory significance. My aim in this paper is not only to contend that their argument fails to show that in confirmation 'there is nothing special about explanatoriness' (Roche & Sober, 2017: 589), but also to reveal what is special confirmationwise about explanatoriness. I will argue that much of the heavy work in Roche and Sober's argument is done by the dichotomy into which they carve up the philosophical options: either explanatory considerations per se are confirmatorily relevant or they acquire whatever confirmatory significance they may possess only by way of background beliefs and so no differently from non-explanatory considerations. I will argue that this is a false dichotomy: Explanatoriness per se can have confirmatory relevance even while explanatory considerations acquire their confirmatory relevance only through background opinions. Furthermore, there is 'something special' about the background opinions concerning explanations that allow explanatoriness in itself to have confirmatory impact. Unlike other sorts of background, having some background opinions expressly about explanations is indispensable to being an observer at all.

1 | INTRODUCTION

In both science and ordinary life, epistemic agents sometimes appear to regard a hypothesis as deriving some plausibility over its rivals from the fact that the explanations it would give (if it were true) are better than those that its rivals would give (if any of them were true). For instance, Darwin explicitly employs this kind of argument:

> 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. The undulatory theory of light has thus been arrived at; and the belief in the revolution of the earth on its own axis was until lately supported by hardly any direct evidence. (Darwin, 1889: 421)

Following Harman (1965), philosophers have termed this 'method of arguing' *inference to the best explanation* (IBE).[1]

'IBE' has sometimes been identified with more ambitious ideas than merely that the explanatory benefits that a hypothesis would bring constitute one of its epistemic virtues. Harman (1965), for instance, holds that all non-demonstrative inference is fundamentally IBE; Lycan (2002:417) calls this view 'ferocious explanationism' and says that it is not widely accepted. The label '*inference* to the best explanation' has sometimes suggested that explanatory considerations warrant inferring the best-explaining hypothesis to be true. This formulation of IBE is too strong (since explanatory considerations alone may be insufficient to justify acceptance of the hypothesis) and this formulation is also too weak (since IBE's advocates regard explanatory considerations as able to support a hypothesis to some degree even before the evidence suffices to justify accepting it). Likewise, IBE has sometimes been identified with the view that the hypothesis H that would give the 'best explanation' of the evidence E should be deemed the most credible or, at least (in Bayesian terms), should be deemed to have the highest prior probability pr(H) or the highest 'likelihood' pr($E|H$) among the rivals (Okasha, 2000: 705). This formulation is also too strong; other considerations can override explanatoriness.

The thesis that I (following others[2]) have just identified with IBE (namely, that the explanatory quality of a hypothesis is sometimes an epistemic reason for according it greater credence than it would otherwise receive) has had notable critics, including Achinstein (2013: 114), Salmon (2001: 74), and van Fraassen (1980: 87; 1989: 160–61). Recently in a series of papers, Roche and Sober (2013; 2014; 2017; 2019) have offered a new argument that explanatoriness lacks confirmatory significance. My aim here is to contest their important argument. I will argue that it fails to show that in confirmation 'there is nothing special about explanatoriness' (Roche & Sober 2017: 589). More importantly, I will identify precisely what is special about the role played by explanatoriness in confirmation.

In section 2, I will present Roche and Sober's argument that explanatoriness is confirmatorily irrelevant. I will argue that much of the heavy work in their argument is done simply by

[1] Lycan (2002) and Douven (2017) survey philosophical work on IBE.

[2] Douven (2017), for example, makes the same identification: 'Abduction, in whatever version, assigns a confirmation-theoretic role to explanation: explanatory considerations contribute to making some hypotheses more credible, and others less so.'

the dichotomy into which Roche and Sober carve up the philosophical options: either explanatory considerations per se are confirmatorily relevant or they acquire whatever confirmatory significance they may possess only by way of background beliefs and so no differently from non-explanatory considerations. In section 3, I will argue that irreducibly explanatory considerations actually play an important confirmatory role in the key example that Roche and Sober use to argue that explanatoriness is confirmatorily irrelevant. This leads to my paper's main point: that Roche and Sober's distinction (between explanatoriness being confirmatory in itself or only by the grace of background opinions) is a false dichotomy. Explanatoriness per se can have confirmatory relevance even while explanatory considerations owe their confirmatory relevance to background opinions. Finally, in section 4, I will argue that Roche and Sober are incorrect in maintaining that 'there is nothing special' about the background opinions concerning explanations that allow explanatoriness in itself to have confirmatory impact. Unlike other sorts of background, having some background opinions expressly about explanations is indispensable to being an observer at all. Section 5 concludes with an overview of my argument and its principal general lessons.

2 | ROCHE AND SOBER'S DICHOTOMY

Here is the key example used by Roche and Sober (2013: 660–661; 2014: 193–94; 2017: 582–83) to argue that explanatoriness is confirmatorily irrelevant. Let S be that Joe was a heavy smoker before age 50 and let L be that Joe contracted lung cancer after age 50. Suppose an agent examines a large random sample of people older than age 50 and finds that in the sample, 30% were heavy smokers before age 50 and 70% of the people who contracted lung cancer after age 50 were heavy smokers before age 50. Joe was not in the sample but is older than age 50. The agent should use her sample frequency information to guide her degrees of confidence $\mathrm{cr}(S)$ and $\mathrm{cr}(S|L)$: suppose that her $\mathrm{cr}(S) = \alpha$ and $\mathrm{cr}(S|L) = \beta$ where $0 < \alpha < \beta < 1$. Roche and Sober (2013: 661) say that it is not essential to their argument that $\alpha = 0.3$ and $\beta = 0.7$, but merely that since the rational agent takes Joe to be 'a random member of the population', she will approximately 'align' her credences regarding Joe with her estimates of the frequencies in the general over-50 population, and the frequencies in the sample are good estimates of the frequencies in that population (2017: 585–86). Roche and Sober then maintain that the further discovery of information E about explanation—namely, that if S and L obtained, then S would explain L—has no confirmatory impact on S: $\mathrm{cr}(S|L) = \mathrm{cr}(S|L\&E)$. That is, the frequency data underwrite the estimate $\mathrm{cr}(S|L)$ of heavy smokers among lung-cancer victims 'and adding the claim that heavy smoking is explanatory doesn't change what that estimate should be' (2013: 661). That the background knowledge of sample frequencies screens off E from S is 'a good explication of the evidential irrelevance of explanatoriness' (2014: 198).

Some aspects of Roche and Sober's argument have been criticized (Climenhaga, 2017; McCain & Poston, 2014; Lange, 2017), and to some of these criticisms, Roche and Sober (2014; 2017; 2019) have replied. But my concern here is not with any of these details, but with Roche and Sober's overall strategy, which has not been critically examined. How is E's confirmatory irrelevance in this example supposed to illuminate the general confirmatory insignificance of facts expressly about explanations?

Here is one way to get ahold of Roche and Sober's strategy. Although in their example above, E is rendered confirmatorily irrelevant by background frequency beliefs, Roche and Sober (2017: 586) allow that 'background assumptions can be cooked up to render E evidentially relevant'. For example, 'if your background information includes the proposition that *~E v Joe is a member of P**, where *P** is a sub-population of people who get lung cancer after age 50, and your estimated

frequency of people who were heavy smokers before age 50 among members of P^* is greater than β, then your credence in the proposition that Joe was a heavy smoker before age 50 should increase to a value greater than β upon learning E' (2017: 588–89). Through this brute-force maneuver (Roche and Sober say), E can be rendered confirmatorily relevant to S, whereas in their original Joe example, E was confirmatorily irrelevant to S. How do Roche and Sober combine these two cases to argue that explanatory considerations have no confirmatory role?

Roche and Sober's aim is not merely to argue that there exist cases (such as their original Joe example) where explanatory considerations (such as E) have no confirmatory impact. Rather, Roche and Sober's point is that even when explanatory considerations have confirmatory significance (as in their second Joe example), explanatory considerations acquire their significance in precisely the same way as non-explanatory facts do: by the grace of background credences that are themselves warranted empirically rather than a priori.

In other words, what Roche and Sober aim to show is that explanatory considerations are no different from other facts in depending entirely on background beliefs for whatever confirmatory significance they may possess in a given case. The confirmatory significance of explanatory considerations is not attributable to some special role for explanation per se that is built into the logic of confirmation—a role that only explanatory considerations can play and that explanatory considerations play purely by virtue of being *explanatory* considerations. There is no fundamental principle of confirmation that gives explanatory considerations some special role. Rather, any principle of confirmation under which a given explanatory consideration has some confirmatory significance derives from more fundamental principles of confirmation that do not privilege explanatory considerations. An explanatory consideration acquires its confirmatory significance in a given case entirely by virtue of possessing some property that non-explanatory considerations sometimes possess (and from which they then derive the same confirmatory significance). In short, Roche and Sober aim to show that 'explanatoriness' has no confirmatory significance *in itself*.

That this is their main point is evident in remarks such as these:

> Our … point is that there is nothing special about explanatoriness: it is like any other contingent proposition in that whether it is evidentially relevant hinges on background information. If explanatoriness is evidentially relevant in a given case, then this is *not* because explanatoriness is evidentially relevant *in itself*. (2017: 589)

> It may seem odd that we have proposed [E's being screened off in the original Joe example] as a good explication of the evidential irrelevance of explanatoriness. After all, this morning's barometric pressure screens-off the barometer reading from a storm this afternoon, and yet the barometer reading is evidentially relevant to there being a storm. Our reply is that the theory of inference to the best explanation is supposed to provide a fundamental epistemology. The idea is that explanatoriness is evidentially relevant in itself; the claim is not that explanatoriness is sometimes correlated with other, more fundamental, properties of a hypothesis that are doing all the epistemic work. (2014: 198)

Although (as we just saw) E is confirmatory in Roche and Sober's second Joe example (unlike in their first), that example nevertheless supports Roche and Sober's main point. That is because the means by which Roche and Sober have 'cooked up' a background that makes E

confirmatorily relevant to *S* would make confirmatorily relevant any arbitrary claim *F* put in *E*'s place (where this place includes having non-zero prior probability and being consistent with *S* together with background knowledge), thereby exemplifying that in confirmation, 'there is nothing special about explanatoriness.' That is, just as the background belief *~E v Joe is a member of P** would enable *E* to confirm *S* despite the background frequency beliefs, so likewise the background belief *~F v Joe is a member of P** would enable an arbitrary *F* to confirm *S* despite the background frequency beliefs. Thus, in the second Joe example, *E* acquires its confirmatory significance only derivatively (through background beliefs), not at all from explanation *in itself* having some special confirmatory significance.

My focus will be on Roche and Sober's dichotomy between explanatory considerations per se being confirmatorily relevant, on the one hand, and their acquiring confirmatory significance only by way of background beliefs and so no differently from non-explanatory considerations, on the other hand. This is the distinction that I will call into question in the next section. For now, though, I want to emphasize merely that it is from this distinction (much more than from their Joe example) that Roche and Sober argue that explanatory considerations play no special role in confirmation.

At least, little more than this distinction is needed to launch such an argument in the form of a dilemma: If explanatory considerations have an impact on confirmation, then either they do so as explanatory considerations per se or through something else. Many philosophers —such as Lipton (2001), McGrew (2003), Okasha (2000), and Sober (1990)—have grasped the latter horn, suggesting that explanatory considerations have their confirmatory impact by making a difference to prior probabilities or likelihoods in Bayesian conditionalization. But then there is 'nothing special' about how explanatory considerations enter into confirmation: they function like any other background. On the other hand, for explanatory considerations to have confirmatory relevance *in themselves* is implausible and arbitrary. Bayesian conditionalization assigns them no special role; such a role would require some fundamental non-Bayesian principle of confirmation expressly privileging them. Some of IBE's friends, such as Lycan (1988; 2012), have posited exactly such an epistemic principle.[3] But to make explanatory potential matter to confirmation by fiat, simply by building it expressly into the notion of a 'good reason' for belief, seems arbitrary. It would leave unanswered the question 'Why not use *schmeason* instead of reason?', where 'schmeason' does not give any special role to explanation.[4]

To posit that explanatory considerations per se are built expressly into good reasonhood seems as unsatisfying as the now widely discredited 'ordinary language dissolution' of the classical problem of induction.[5] According to that approach, part of what we *mean* by evidence rendering a given scientific theory 'justified' is simply that there is a strong inductive argument for the theory from the evidence. This approach tells us nothing about why we ought to adopt 'justified' beliefs.[6]

[3] Lycan (1988: 134) characterizes epistemic principles privileging explanatory considerations as 'ultimate, not themselves justified by any more fundamental epistemic norms'; he offers a philosophical justification (by reflective equilibrium with our epistemic intuitions) of this conception of IBE. Van Fraassen's (1989: 160–61) interpretation of IBE as awarding 'bonus points' to hypotheses in recognition of their explanatory potential is one approach that such a friend of IBE might adopt (though IBE's friend must then defend IBE against van Fraassen's critique of IBE).

[4] Sober (2015a: 43): 'If reason, by definition, obliges us to have these goals [of finding theories that form a certain sort of parsimonious, unified system], why shouldn't we say "no" to reason and sign up under the banner of *schmeason*, which requires no such commitment?'

[5] Defenders of the ordinary-language approach include Barker (1974), Horwich (1982: 97–98), and Strawson (1952: 256–57); critics include BonJour (1998: 196–99), Salmon (1957; 1967: 49–52), and Skyrms (1986: 47–54).

[6] Salmon (1957: 42): 'It sounds very much as if the whole [ordinary-language] argument has the function of transferring to the word "inductive" all of the honorific connotations of the word "reasonable" quite apart from whether induction is

Likewise, for explanatoriness to matter in itself would be for its mattering to rest on no more fundamental principle and so would leave nothing to be a reason why explanatoriness matters. But if explanatoriness's confirmatory significance is not to be ad hoc, then its significance must fall out of some broader account of confirmation. We are then thrust back onto the dilemma's other horn: when explanatory considerations matter, they do not matter in themselves, but only by virtue of possessing some property that non-explanatory considerations sometimes possess, thereby making those considerations matter, and so 'there is nothing special' confirmationwise about explanatoriness.

Much of Roche and Sober's argument, then, derives simply from their carving up the terrain in terms of the dichotomy between explanatoriness being confirmatory in itself or only by the grace of empirical background beliefs. Let's now see how Roche and Sober's Joe example also depends on this dichotomy. We will then be able to see how that dichotomy misleads.

3 | THE FALSE DICHOTOMY

Roche and Sober regard explanatory considerations as having no confirmatory impact in their first Joe example. I will now suggest, by contrast, that although Roche and Sober are correct that *E*'s discovery (after the discovery of the random sample's frequencies) has no confirmatory impact, irreducibly explanatory considerations nevertheless play an important (though hidden) role in their example. In particular, the agent in their example presumes that the *frequency* of heavy smokers in the population from which the sample was randomly drawn is a good reflection both of the *chances* that Joe encounters various causes of smoking (whatever the causes of smoking may be) and of the *powers* of those causes, if Joe encounters them, to promote or to inhibit Joe's smoking.

This presupposition about the factors eligible to explain Joe's smoking (if he smokes) is crucial to the relevance of the sample's frequency in guiding the agent's credences regarding Joe. For its relevance diminishes or disappears entirely when the agent does not make this presupposition (and no other suitable prior opinion takes its place). For example, suppose that the agent makes no such presupposition because the agent's background beliefs include that Joe is not a human being but is instead an extraterrestrial living in a distant galaxy (whereas the population from which the sample was randomly drawn consists entirely of our terrestrial human contemporaries). Then, presumably, the agent believes that the reasons why Joe does (or does not) smoke are utterly different from the reasons why the individuals in the sampled population do (or do not). In the absence of some further presupposition, the agent regards the frequency of smokers in the general over-50 population (or in its sub-population with lung-cancer)—or in the sample drawn randomly from there—as no guide to her credence in *S* (or in *S* given *L*), which exclusively concern the extraterrestrial Joe.

Here is another example illustrating the confirmatory relevance of the agent's presupposition about explanation. (Henceforth let's assume that the agent knows that we are dealing exclusively with human beings!) Rather than the agent's considering the frequency of smokers in the general over-50 population (from which the agent draws a random sample), suppose the agent deliberately gerrymanders a sub-population by combining Joe with various individuals all of whom the agent knows to have smoked heavily before age 50. Obviously, the agent knows the frequency of heavy smokers in this sub-population is very high, since the agent knows that at least all but

good for anything. The resulting justification of induction amounts to this: If you use inductive procedures you can call yourself "reasonable"—*and isn't that nice!*'

one of its members are heavy smokers. (To ascertain this fact, the agent had no need to draw a random sample from this sub-population.) But this high frequency is not a good guide for the agent in arriving at her credence in *S*. Why does the agent have no reason to align her credence in *S* with this frequency but good reason to align it with the frequency of heavy smokers in the general over-50 population (as indicated by the frequency in the sample drawn randomly from that population)? Because of the differences in the agent's background knowledge regarding the two populations: that the frequency in the general population is indicative of the prevalence and strength of the potential causes and hence explainers of Joe's smoking (or not smoking), whereas the frequency in the gerrymandered sub-population is not.

In Roche and Sober's Joe example, then, we cannot conclude that explanatory considerations are confirmatorily irrelevant merely from *E*'s confirmatory irrelevance to *S*. Some explanatory considerations have their confirmatory impact upstream from *E*'s discovery—namely, in making the frequencies (in the sample or in the population sampled) relevant to the agent's credences regarding Joe.

In Roche and Sober's example, the confirmatory role of the agent's background explanatory knowledge is obscured by the power of the agent's frequency knowledge to render *E* confirmatorily irrelevant. No explanatory considerations appear to be at work—until we ask about the rationale for the frequencies' own relevance to the agent's credence in *S*. But in other examples, there are no known frequencies to obscure the role of explanatory background opinions in making the evidence relevant to the hypothesis it confirms. For example, before the laws of electrostatics were discovered, it was believed (on the basis of an experiment by Benjamin Franklin) that a charged, electrically conductive, hollow shell exerts no electrostatic force on a charged body anywhere in its hollow interior.[7] It was known that this null result follows mathematically from a hypothetical electrostatic force law where the force between point charges is proportional to the inverse-square of their separation, just as Newton's inverse-square gravitational-force law entails and explains why a mass inside a spherical shell of uniform density feels no gravitational force from the shell.[8] The fact that gravity is governed by an inverse-square force was then regarded (by Joseph Priestley, among others) as some evidence that the electrostatic force is too.[9] What made gravity so confirmatorily relevant to electrostatics? That they are alike in displaying the null result. In particular, that the two forces have this null result in common was considered some evidence that the two forces (and hence the two null results) have the same sort of explanation, so the fact that the explanation of gravity's null result is an inverse-square law was some evidence that the explanation of electrostatics' null result is an inverse-square law. In this example, there is no random sampling from a large population of kinds of forces (exhibiting null results), and so there is no frequency with which the randomly sampled forces obey inverse-square laws. So there is no knowledge of frequencies to render explanatory facts confirmatorily irrelevant.

[7] Franklin's experiment concerned a charged metal cup, not a spherical shell (Heilbron 1979: 464). I will ignore this complication since the scientists at the time apparently saw it as no obstacle to the experiment's relevance.

[8] It was *not* known whether an inverse-square law is the only mathematically simple function of distance that as a force function is capable of yielding the null result by integration over the contributions to the total force exerted by of the various shell elements. (See Heilbron 1979: 464.)

[9] Priestley wrote in 1767 (Priestley 1794: 640, quoted in Heilbron 1979: 464): 'May we not infer from this experiment, that the attraction of electricity is subject to the same laws with that of gravity, and is therefore according to the squares of the distances; since it is easily demonstrated, that were the earth in the form of a shell, a body in the inside of it would not be attracted to one side more than another.' There is no need to maintain that the evidence here is decisive (as Priestley's remark suggests), merely that it confirms the electrostatic inverse-square law. But I will say more in a moment about what precisely 'the evidence' here is.

Of course, no explanatory background is needed to make the null electrostatic result confirm that the electrostatic force law is inverse-square. It suffices for this confirmation (given that the prior probabilities of the inverse-square electrostatic law and the null electrostatic result were neither 0 nor 1) that the hypothesis of an inverse-square electrostatic law entails the null electrostatic result. But the question is not why the null electrostatic result confirms that the electrostatic force law is inverse-square, but rather why—even on top of that confirmation—the inverse-square character of the law governing the *gravitational* force (which exhibits a null result) further confirms that the law governing the electrostatic force (which also exhibits a null result) is inverse-square. (Obviously, there is no logical entailment between the gravitational-force law's being inverse-square and anything about electrostatics.) The connection here between the confirming evidence and the hypothesis being confirmed (the connection that underwrites the confirmation) is nothing less than an analogy between the two *explanations* (one known and the other being confirmed): both would explain null results, and the two laws would explain these analogous results in the same way by virtue of possessing the same mathematical property (namely, being inverse-square). Nothing less than this analogy between the two explanations—no fact solely about logical entailments, for instance—suffices to capture the connection that powers the confirmation. That an inverse-square electrostatic law is what explains the electrostatic null result is supported (to some degree, at least) by the fact that a similar phenomenon (the gravitational null result) is explained by the inverse-square character of another law.[10]

The confirmation in this historical episode depends on a presumed analogy between the two explanations. But rationality alone fails to require an agent to make this presumption. This opens the door to an objection to my foregoing argument. This objection invokes the dichotomy from Roche and Sober that I highlighted in the previous section: the dichotomy between explanatoriness in itself being confirmatorily relevant, on the one hand, and explanatoriness owing its confirmatory relevance to background opinions, on the other hand. Roche and Sober could use this distinction to object that at best, my foregoing argument shows merely that explanatory considerations play a role in confirmation only when background beliefs enable them to do so and therefore function no differently confirmationwise from any other discovery. My argument fails to show that explanatory considerations are confirmatorily relevant in themselves. For instance, that the gravitational null result is explained by an inverse-square law confirms that the electrostatic null result is, too, only by the grace of some significant degree of background confidence that two such analogous phenomena have analogous explanations. Such confidence (the objection continues) is no different from the background belief (mentioned in the previous section) that *~E v Joe is a member of P**, which would allow *E* to confirm *S*. My argument reveals 'nothing special' about how explanatoriness acquires its confirmatory impact.

My reply to this objection takes me to one of my principal points: that Roche and Sober's alleged distinction between explanatory considerations being confirmatory in themselves or only by the grace of background opinions is a false dichotomy. Explanatory considerations per se can have confirmatory relevance even while explanatory considerations depend on background opinions for their confirmatory relevance. The background opinions that give explanatory considerations their confirmatory significance can themselves be about explanations per se (rather than about logical entailments or frequencies or anything other than explanations). In that case, discoveries of explanations matter confirmationwise in themselves, not because explanatoriness is

[10] That Priestley himself regarded this as an argument by analogy is supported by a passage from Priestley that Francesco Nappo called to my attention (in the course of his work on arguments by analogy): 'analogy is our best guide in all philosophical investigations' (Priestley 1794: 380).

associated with something else—and yet these discoveries owe their confirmatory power to these background opinions about contingent matters of fact (namely, about explanations).

Thus, explanatory considerations do not need to possess their confirmatory relevance independent of background opinions about contingent matters of fact in order for explanatoriness per se to have confirmatory relevance—as long as those background opinions are also about explanatory considerations per se. This moral is illustrated by the electrostatics case. The confirming evidence (that an inverse-square law *explains* the gravitational null result) expressly concerns explanation. It is thereby connected to background opinions about explanations per se (namely, some non-negligible confidence that the gravitational and electrostatic null results have similar *explanations*). It is the evidence's concern with explanation that enables the evidence to confirm the electrostatic inverse-square law as much as it does.

Roche and Sober might reply that when *E* confirms *S* in light of the background *∼E v Joe is a member of P**, it is the evidence's concern with explanation (in particular, that the evidence is *E*) that enables the evidence to confirm *S*. Such a case (Roche and Sober might point out) nevertheless fails to involve explanatory considerations per se having confirmatory relevance. I would agree: that example was just 'cooked up' (as Roche and Sober say) by using a logical trick to render *E* confirmatorily relevant. As we saw, *E* could be systematically replaced in the evidence and background by an arbitrary proposition *F*, even one not concerning explanation, and the same maneuver would then make *F*'s discovery confirm *S*.

But the same point cannot be made when the background is some confidence that the gravitational and electrostatic null results have similar explanations and the evidence is the inverse-square explanation of the gravitational null result. Admittedly, from the background and evidence we could systematically remove all references to the gravitational null result and replace them with references to some other, arbitrary phenomenon. Background confidence that this phenomenon and the electrostatic null result have similar explanations might then enable this other phenomenon's explanation to confirm some hypothesis about the electrostatic null result's explanation. But this confirmation would still involve the confirmatory relevance of explanatoriness per se – that is, would still involve explanatory considerations having their confirmatory significance by virtue of their expressly concerning explanation.

Admittedly, once again, rationality alone fails to require an agent to have significant confidence that the gravitational and electrostatic null results have similar explanations, just as rationality alone fails to require an agent to believe *∼E v Joe is a member of P**. Nevertheless, although no *particular* background opinion about explanations is required by rationality alone, I will argue (in the next section) that any agent who is in a position to consider whether some new empirical evidence confirms a given hypothesis is rationally required to have significant justified confidence in *many* hypotheses expressly concerning the explanations of various facts—and also in *many* hypotheses that various cases have similar (though perhaps as yet unknown) explanations. There is no analogous requirement regarding some class of hypotheses like *∼E v Joe is a member of P**. So there is indeed 'something special' about explanatory considerations: we are justly able to bring them to bear in confirmation from the dawn of our epistemic lives. They are therefore available to justly influence our confidence in various potential explanations even in cases where very little other relevant information is available to us.

4 | THE INDISPENSABILITY OF HAVING EXPLANATORY BACKGROUND OPINIONS

The account that I will now give inevitably presupposes answers to some longstanding fundamental epistemological questions. I cannot argue fully for those answers here. Accordingly, what I will now give is perhaps less an argument than a how-possibly explanation: one possible way for background opinions about explanations to be indispensable to observers and hence one possible way for there to be 'something special' about explanatory considerations. (Accordingly, this proposal does not preclude other ways, given other epistemological views, for observers unavoidably to have some large store of background opinions regarding explanations per se.)

A piece of direct perceptual knowledge (such as my knowledge that my dog is currently sitting here next to me, which I see is the case) is not acquired as the conclusion of an inference and does not acquire its status as justified by inheriting it inferentially. But in order for an observer's perceptual belief to acquire its justificatory status non-inferentially, she must recognize somehow that her perceptual belief acquires its justificatory status from its having a reliable source, namely, her own dispositions to respond perceptually in various ways to her surroundings. That is, she must recognize that considering her perceptual belief's etiology, what she purports to have observed is likely to have been the case (absent any information to the contrary). This recognition requires that she be in a position to infer the (likely) truth of her purported observation from her having made it and from the conditions under which she did so. This internalist requirement applies whether the observation that she purports to have made is one that the vast majority of competent language-users are able to make reliably under her conditions or whether it instead requires specialized training undergone by few language-users.

DeVries (2005: 111–12) terms this requirement 'Epistemic Reflexivity' (ER):

> To know some particular fact through the exercise of some cognitive capacity C, one must know that C is a reliable source of information about such facts under the circumstances in which one knows that C is being exercised in this case.

ER's advocates include BonJour (1985: 116), Rosenberg (2002; 2007), Sellars (1963a: 168–70, §§35–38; 1975: 325–26, ii §§40–41), and Williams (2001: 175–79; 2009: 164, 176).[11] One motivation for ER is that an observer should treat her own perceptual capacities as she should treat someone else's and as she should treat some object's non-perceptual responsive dispositions. For any of them to be one's source of knowledge about the responder's environment, one must be justly confident that the responder's responses are reliable indicators of its environment. This requirement applies when the response is a litmus paper's turning red in response to surrounding acid, a well-trained radiologist's saying 'I see a broken leg' in response to an X-ray image, or an ordinary language-user's saying 'I see a red apple.' ER applies this requirement to one's own perceptual responses.

ER is also motivated by the idea that an observer should treat her own observations as she should treat any knowledge that acquires its justificatory status at least partly in virtue of its origin (e.g., an expert's testimony). In any such case, the knower must be in a position to infer from the report's origin to its truth. Sellars (1963b: 88) calls this a 'trans-level inference' because relative to

[11] Notable critical discussions of these defenses of ER include Alston (1989), Brandom (1994: 218–21; 1997: 157–59), and Sosa (1997). One important concern is that ER requires so much of an observer that it excludes infants and animals from qualifying as observers. I cannot address this concern here, though one move in the right direction may be to exploit the difference between being in a position to make a given inference and having actually made it (see the following note).

its conclusion (*p*), some of its premises are at a meta-level (since they include the agent's having come to believe that *p* in certain circumstances).

Another motivation for ER (Rosenberg 2002: 116–117; Williams 2009: 176–80) is that an observer must be prepared to disavow her purported observation if she discovers that the responsive disposition she exercised in making her report had been unreliable under the conditions in which she made that report. If she is not so prepared, then she cannot be held responsible for her response; it was just a reflex, having no justificatory status.

According to ER's advocates, an observer's knowledge of the reliability of one of her own responsive dispositions (under certain conditions) is acquired *inductively* (Sellars 1963a: 169, §37). As she accumulates a track record of responding accurately (usually) to some particular sort of fact, she is entitled to increasing confidence in her disposition's reliability. Eventually, she has sufficient evidence to be able to infer justly, from her purportedly observing that (e.g.) there is a red apple before her, that there is indeed such an apple.[12]

Which episodes in her track record confirm the responder's accuracy in a given case where she purportedly observed that there is a red apple before her? Not every instance where she says that she sees a red apple. In some instances, the conditions are known to be so different from those prevailing in the given case that her track record in those instances should be considered irrelevant to the given case. Furthermore, some episodes where she reports seeing things that are *not* red apples (such as other red things or other kinds of fruit or other colors) may well be confirmatorily relevant to her reliability in visually identifying red apples under certain environmental conditions, especially if those conditions are the same as those prevailing in all of the episodes (e.g., in sunlight). Indeed, a well-trained physician may have no track record at all of purportedly observing a given rare circumstance (e.g., having never personally encountered a case of the rare genetic disorder Cornelia de Lange Syndrome), but her extensive track record of identifying on sight various other, related, more common circumstances (such as other genetic disorders) may confirm her reliability in observing the rare circumstance. In addition, the track records of *other* responders may suffice (in the absence of contrary evidence) to strongly confirm, of any mature language-user, that she can reliably observe a given ordinary sort of fact.

What makes all of these diverse pieces of evidence relevant confirmationwise to the responder's accuracy in the given case? Explanatory considerations. Insofar as some past episode is believed to involve the exercise of the same responsive disposition (i.e., the same sort of perceptual mechanism) in the same sort of environment as the given episode involves, to that degree a responder's accuracy in the past episode confirms the given responder's accuracy in the given case. The past episode and the given case need not be believed to involve exactly the same causal mechanism;

[12] There need not be some sharp moment at which her evidence becomes strong enough to justify the inference. How strong it must be may be vague and context sensitive; in this respect, this induction is like many others. The observer must be in a position to make this inference but need not ever have explicitly made it. Indeed, Rosenberg (2002: 117–118) notes that an observer called upon to justify her confidence in her responsive disposition's reliability would typically offer something less formal than (but able to be elaborated as) an induction from her track record, such as 'It's broad daylight,' 'I'm not colorblind, you know,' or 'I know a red apple when I see one.' Sellars (1963a: 168–69, §§36–37) emphasizes that her successful responses early in her track record were not *at the time* observations since she did not then know that her disposition was reliable. But once her track record becomes longer, she may accumulate evidence that justifies believing that it was reliable *at that earlier moment*. That earlier episode then becomes an observation retroactively. This change is not backwards causation because 'in characterizing an episode or a state as that of *knowing*, we are not giving an empirical description of that episode or state; we are placing it in the logical space of reasons' (Sellars, 1963a: 169, §36). See Rosenberg (2002) and Williams (2001: 176–78; 2009) for complementary attempts to grapple with ER's threat of vicious epistemic circularity (from observations requiring knowledge of reliability that, in turn, is acquired from a track record of events that were physically but not yet normatively just like observations).

the degree of confirmation reflects the extent of the mechanisms' purported similarity. In this way, the reliability of the responder's disposition to respond to green pears (in sunlight) can be confirmatorily relevant to the reliability of her disposition to respond to red apples (in sunlight). When the given disposition is believed to be widely shared, a given responder's reliability can be confirmed by the track records of other responders under similar conditions, since they are believed to be engaging the same mechanism.

Of course, an agent could have no beliefs about the *details* of the mechanisms producing various responses and still have justified confidence that the mechanism (whatever it is) that produced one response is similar to the mechanism that produced another. But if she has no opinions at all about which episodes involve the exercise of similar responsive dispositions and so have similar (though perhaps unknown) explanations, then (in the absence of other information) she has no basis for demarcating what track record is confirmatorily relevant—and therefore (by ER) she is no observer. To treat your responding 'I see a green pear' (in sunlight) but not my responding 'I see a red apple' (under sodium-arc lamplight) as bearing on my reliability in identifying red apples (in sunlight) makes sense in many instances only when understood as guided by explanatory considerations.[13]

Here is another way to appreciate the role played by background opinions about explanations in making the evidence of someone's past record of putative observations able to confirm (or disconfirm) her accuracy in responding to some new case. Let's turn from her track record's confirmatory relevance at a given moment to the way that its relevance changes with the discovery of additional facts about the episodes in her track record. The range of considerations that can enhance, diminish, or entirely eliminate a track-record episode's confirmatory relevance to the truth of a given purported observation (e.g., that the area was illuminated by sodium-arc lamplight, that the responder took a given psychedelic drug) is essentially open-ended. How can one have such an open-ended commitment now to modifying the degree to which the track record confirms a given purported observation? By taking this confirmation to depend upon explanatory background information. One can be committed now to aligning a track-record episode's confirmatory relevance with whatever one's opinion later may be regarding how the mechanism exercised in that episode compares to the mechanism exercised in the purported observation being confirmed to be accurate.[14] In this way, one can undertake a commitment now regarding the way that future discoveries *that cannot yet even be formulated* would influence the track record's confirmatory relevance.

Thus, explanatory considerations are indispensable guides for the inductive arguments that (by ER) observers must be capable of making.[15] It might be objected that an observer need not have

[13] Consider the argument that one would not be justified in taking certain past episodes in my track record, but not others, as bearing confirmationwise on my reliability in making a putative observation in a given case, unless in doing so one were guided by background opinions about which episodes involve the exercise of dispositions explained by similar mechanisms. This argument is like Sellars's argument (1963b: 80–87; 1963c: 118–26, §§33–54; 1977: 319) that science would not be justified in taking certain empirical generalizations as well-confirmed by our observations unless science were guided by certain explanatory considerations—in particular, unless science were operating in a framework positing certain unobservables. (For example, those generalizations would appear gerrymandered and unmotivated in a phenomenalist framework.)

[14] In Roche and Sober's Joe example, we saw that explanatory considerations influence the sample frequency's confirmatory relevance. Likewise, in the track-record confirmation, explanatory considerations influence the confirmatory relevance of the frequencies with which various classes of purported observations have been accurate.

[15] As I mentioned at this section's start, that observers must be prepared to make such inductive arguments does not entail that observation reports are the products of inferences or that their justification is inferential.

any opinions regarding the explanations of the episodes in what I have called 'her track record'. The objector might insist that an observer can justly use the accuracy of her past observation reports to confirm the accuracy of her current purported observations without appealing to any explanatory considerations, even though in the absence of explanatory considerations, the range of past episodes that she takes to confirm the accuracy of a given purported observation appears arbitrary or gerrymandered (as I suggested in connection with my reliability in identifying the green pears in sunlight, but not in identifying the red apples under sodium-arc light, as bearing on my accuracy in identifying the red apples in sunlight).

To this objection, I would reply that an observer cannot avoid trafficking in explanations. That is because an observation report's content is not merely (e.g.) 'There is a red apple before me,' but rather '*I see that* there is a red apple before me.' That is, the observation report claims that a given circumstance (that there is a red apple before me) not only obtains but also helps to explain why the reporter made the report: because that circumstance engaged the responsive disposition that produced the purported observation. Observing is simultaneously explaining.[16] As Rosenberg (2002: 242, cf. 88, 92) puts it: 'a spontaneous perceptual judgement already expresses a first attempt to account for its own occurrence by incorporating an explanatory belief regarding its origins; that is, a hypothesis regarding the character of the item which, *qua* stimulus, has (causally) *evoked* it.' Observations therefore inevitably presuppose some background explanatory opinions.[17]

The inference from a responder's purportedly making a particular observation to the observation report's (likely) truth is, then, a case of IBE. The inference's conclusion is that the fact purportedly observed to be the case explains the observation's having been made. The inference's strength is modulated by a background of expressly explanatory considerations— namely, that the track record consists of episodes purportedly having explanations similar to the purported observation's explanation. Background opinions about explanations per se are needed to carve out the range of episodes that are confirmatorily relevant to the conclusion that an observer must be prepared to draw.

There is thus 'something special' about background beliefs about explanations. Unlike other sorts of background, having some such background is indispensable to being an observer at all. It is not open to an observer to do without any expressly explanatory background opinions; she cannot use 'schmeason' (which fails to attend to any explanatory considerations per se in confirming hypotheses) rather than reason.

5 | CONCLUSION

Let me summarize the path that we have taken. According to IBE, the explanatory quality of a hypothesis is sometimes an epistemic reason for according it greater credence. Roche and Sober

[16] Such a view might take inspiration from Peirce's remarks that 'abductive inference shades into perceptual judgments without any sharp line of demarcation between them … our first premisses, the perceptual judgments are to be regarded as an extreme case of abductive inferences' (Peirce, 1932: 224) and 'nothing is more familiar (especially to every psychology student) as the interpretativeness of the perceptive judgment. It is plainly nothing but the extremest case of Abductive Judgment' (Peirce, 1932: 229). But Peirce's work on perception is terribly obscure, and his 'abduction' is not quite the same as IBE.

[17] That perceptual evidence requires some background opinions regarding explanations is not well-reflected in the standard Bayesian notation pr(H|E,B) suggesting that E and B are independent, i.e., that B could be altered without losing E.

have recently posed a dilemma for IBE: either explanatory considerations per se are confirmatorily relevant or they acquire whatever confirmatory significance they possess only by the grace of background beliefs. I have argued that we can grasp both horns of this apparent dilemma: Explanatoriness per se can have confirmatory relevance even while explanatory considerations acquire their confirmatory significance only through background opinions. That is because the background opinions making a hypothesis's explanatory qualities confirmatorily relevant can also be about explanations per se.

On this view, the explanatory quality of some hypothesis has its confirmatory significance only because some background opinions make it confirmatorily significant. Therefore, explanatory quality acquires its confirmatory significance no differently from the way that other considerations do. This result seems to show that (as Roche and Sober put it) 'there is nothing special' about IBE. IBE's foes have often argued against IBE by taking some plausible account of confirmation and pointing out that it accords no special role to explanation (e.g., Sober, 2015*b*: 913). IBE is set apart from other views of confirmation by according explanation per se a distinctive epistemic role.

However, I have argued that although explanatory considerations have an impact on confirmation only when background beliefs enable them to do so (and in this respect, explanatory considerations function no differently confirmationwise from other considerations), there remains 'something special' about explanatoriness in confirmation. Unlike other sorts of background opinions, an epistemic agent's having some background opinions expressly about explanations is indispensable to the agent's being able to make observations.

Furthermore, each observation that an agent makes gives her not only some justified beliefs[18] concerning some features of her environment, but also some justified beliefs concerning some explanations—namely, concerning what features of her environment *explain why* she responded to her environment as she did (that is, by having the perceptual experience she had and making the observation she made). As a result, an observer is guaranteed to have a great store of justified background beliefs about explanations per se. Therefore, an epistemic agent (whether an ordinary observer or a scientist) will automatically have many justified background beliefs about other explanations to draw upon in confirming a given hypothesis— even when she has very little by way of other relevant justified background beliefs to enable evidence to confirm or to disconfirm that hypothesis.

This constitutes another important respect in which 'there is something special' about IBE. Suppose that a given hypothesis would explain certain known facts in much the same way as other facts are already known to be explained, where we justly expect all of these facts to have similar sorts of explanations. Their similarity to these other known explanations contributes to the quality of the explanations that the given hypothesis would supply. Therefore, by IBE, the hypothesis derives some increased plausibility from these other known explanations. IBE is thus able to play an especially important role in confirmation because it can still operate even when other sorts of background beliefs that it would be useful to possess regarding the case at hand (like the information about frequencies in Roche and Sober's example) are unavailable.

In this way, even if a hypothesis's qualities as an explainer can play a role in confirmation only with the help of suitable background beliefs, IBE can have a distinctive and important role in confirmation.

[18] Or, if the observation is not captured by full belief in the truth of some proposition (as in Jeffrey's (1983) example of someone observing the colors of jellybeans that are illuminated only by dim candlelight), then at least the observation leaves her with justified changes to her credences.

REFERENCES

Achinstein, P. (2013). *Evidence and Method*. Oxford: Oxford University Press.

Alston, W. (1989). What's wrong with immediate knowledge? In W., Alston, *Epistemic Justification* (pp. 67–81). Ithaca: Cornell University Press.

Barker, S. (1974). Is there a problem of induction? In R. Swinburne (Ed.), *The Justification of Induction* (pp. 57–61). London: Oxford University Press.

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

BonJour, L. (1998). *In Defense of Pure Reason*. Cambridge: Cambridge University Press.

Brandom, R. (1994). *Making It Explicit*. Cambridge: Harvard University Press.

Brandom, R. (1997). Study guide. In R. Brandom (Ed.), *Empiricism and the Philosophy of Mind* (pp. 119–81). Cambridge: Harvard University Press.

Climenhaga, N. (2017). How explanation guides confirmation. *Philosophy of Science*, *84*, 359–68. https://doi.org/10.1086/690723

Darwin, C. (1889). *The Origin of Species*, 6th ed. London: Murray.

deVries, W. (2005). *Wilfrid Sellars*. Montreal and Kingston: McGill-Queen's University Press.

Douven, I. (2017). Abduction. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, (Summer 2017 Edition), URL = https://plato.stanford.edu/archives/sum2017/entries/abduction/. Last accessed 26 September 2018.

Harman, G. (1965). The inference to the best explanation. *The Philosophical Review*, *74*, 88–95. https://doi.org/10.2307/2183532

Heilbron, J. L. (1979). *Electricity in the 17th and 18th Centuries*. Berkeley and Los Angeles: University of California Press.

Horwich, P. (1982). *Probability and Evidence*. Cambridge: Cambridge University Press.

Jeffrey, R. C. (1983). *The Logic of Decision*, 2nd ed. Chicago: University of Chicago Press.

Lange, M. (2017). The evidential relevance of explanatoriness: a reply to Roche and Sober. *Analysis*, *77*, 303–312. https://doi.org/10.1093/analys/anx045

Lipton, P. (2001). Is explanation a guide to inference? In G. Hon & S. Rakover (Eds.), *Explanation: Theoretical Approaches and Applications* (pp. 93–120). Dordrecht: Kluwer. https://doi.org/10.1007/978-94-015-9731-9_4

Lycan, W. G. (1988). *Judgement and Justification*. Cambridge: Cambridge University Press.

Lycan, W. G. (2002). Explanation and epistemology. In P. Moser (Ed.), *The Oxford Handbook of Epistemology* (pp. 408–33). Oxford: Oxford University Press. https://doi.org/10.1093/0195130057.003.0015

Lycan, W. G. (2012). Explanationist rebuttals (coherentism defended again). *The Southern Journal of Philosophy*, *50*, 5–20. https://doi.org/10.1111/j.2041-6962.2011.00087.x

McCain, K., & and Poston, T. (2014). Why explanatoriness is evidentially relevant. *Thought*, *3*, 145–53. https://doi.org/10.1002/tht3.122

McGrew, T. (2003). Confirmation, heuristics, and explanatory reasoning. *The British Journal for the Philosophy of Science*, *54*, 553–67. https://doi.org/10.1093/bjps/54.4.553

Okasha, S. (2000). Van Fraassen's critique of inference to the best explanation. *Studies in History and Philosophy of Science*, *31*, 691–710. https://doi.org/10.1016/s0039-3681(00)00016-9

Peirce, C. S. (1932). *The Collected Papers of Charles S. Peirce*, vol. 2. Cambridge: Harvard University Press.

Priestley, J. (1794). *The History and Present State of Electricity, with Original Experiments*, 5th ed. London: J. Johnson and F. C. Rivington.

Roche, W., & Sober, E. (2013). Explanatoriness is evidentially irrelevant, or inference to the best explanation meets Bayesian confirmation theory. *Analysis*, *73*, 659–668. https://doi.org/10.1093/analys/ant079

Roche, W., & Sober, E. (2014). Explanatoriness and evidence: a reply to McCain and Poston. *Thought*, *3*, 193–99. https://doi.org/10.1002/tht3.128

Roche, W., & Sober, E. (2017). Is explanatoriness a guide to confirmation? A reply to Climenhaga. *Journal for General Philosophy of Science*, *48*, 581–90. https://doi.org/10.1007/s10838-016-9357-5

Roche, W., & Sober, E. (2019). Inference to the best explanation and the screening-off challenge. *Teorema*, *38*, 121–42.

Rosenberg, J. (2002). *Thinking about Knowing*. New York: Oxford University Press.

Rosenberg, J. (2007). Sellarsian seeing. In J. Rosenberg, *Wilfrid Sellars: Fusing the Images* (pp. 243–65). New York: Oxford University Press.

Salmon, W. C. (1957). Should we attempt to justify induction? *Philosophical Studies*, *8*, 33–48. https://doi.org/10.1007/bf02308902

Salmon, W. C. (1967). *The Foundations of Scientific Inference*. Pittsburgh: University of Pittsburgh Press.

Salmon, W. C. (2001). Explanation and confirmation: a Bayesian critique of inference to the best explanation. In G. Hon, & S. Rakover (Eds.), *Explanation: Theoretical Approaches and Applications* (pp. 61–92). Dordrecht: Kluwer. https://doi.org/10.1007/978-94-015-9731-9_3

Sellars, W. S. (1963a). Empiricism and the philosophy of mind. In Sellars (1963d), pp. 129–96.

Sellars, W. S. (1963b). Phenomenalism. In Sellars (1963d), pp. 60–105.

Sellars, W. S. (1963c). The language of theories. In Sellars (1963d), pp. 106–126.

Sellars, W. S. (1963d). *Science, Perception and Reality*. London: Routledge & Kegan Paul.

Sellars, W. S. (1975). The structure of knowledge. In H.-N. Castañeda (Ed.), *Action Knowledge and Reality* (pp. 295–346). Indianapolis: Bobbs-Merrill.

Sellars, W. S. (1977). Is scientific realism tenable? In *PSA 1976*, *2*, 307–34. https://doi.org/10.1086/psaprocbienmeetp.1976.2.192388

Skyrms, B. (1986). *Choice and Chance*, 3rd ed. Belmont, CA: Wadsworth.

Sober, E. (1990). Let's razor Ockham's razor. In D. Knowles (Ed.), *Explanation and its Limits* (pp. 73–93). Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511599705.006

Sober, E. (2015*a*). *Ockham's Razors: A User's Manual*. Cambridge: Cambridge University Press.

Sober, E. (2015*b*). Two Cornell realisms: moral and scientific. *Philosophical Studies*, *172*, 905–924. https://doi.org/10.1007/s11098-014-0300-5

Sosa, E. (1997). Mythology of the given. *History of Philosophy Quarterly*, *14*, 275–87.

Strawson, P. F. (1952). *An Introduction to Logical Theory*. London: Methuen.

van Fraassen, B. C. (1980). *The Scientific Image*. Oxford: Clarendon.

van Fraassen, B. C. (1989). *Laws and Symmetry*. Oxford: Clarendon.

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

Williams, M. (2009). The tortoise and the serpent: Sellars on the structure of empirical knowledge. In W. deVries (Ed.), *Empiricism, Perceptual Knowledge, Normativity, and Realism* (pp. 147–85). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199573301.003.0007

How to cite this article: Lange, M. (2024). A false dichotomy in denying explanatoriness any role in confirmation. *Noûs*, *58,* 520–535. https://doi.org/10.1111/nous.12469