

teorema

Vol. XXXIX/2, 2020, pp. 27-42

ISSN: 0210-1602

[BIBLID 0210-1602 (2020) 39:2; pp. 27-42]

What Inference to the Best Explanation Is Not: A Response to Roche and Sober's Screening-Off Challenge to IBE

Marc Lange

RESUMEN

Roche y Sober (2013) han ofrecido un nuevo argumento de neutralización [“screening off” argument] en contra de la inferencia a la mejor explicación [“Inference to the Best Explanation”, “IBE” por sus siglas en inglés]. La forma en que Roche y Sober deben concebir a la IBE para que su argumento sea aplicable queda revelada en su reciente respuesta (2019) a las objeciones de Lange (2017) a su argumento de neutralización. El argumento de Roche y Sober en contra de la IBE requiere que bajo la IBE las “virtudes explicativas” que posee una hipótesis H modifiquen la credibilidad que tiene un agente racional en H independientemente de sus opiniones de fondo. Pero es impropio atribuir a la IBE tal conexión mágica entre las consideraciones explicativas y la confirmación.

PALABRAS CLAVE: *bayesianismo; relevancia evidencial; explicatividad; inferencia a la mejor explicación; Sober; neutralización; unificación.*

ABSTRACT

Roche and Sober (2013) have offered a new, “screening-off” argument against Inference to the Best Explanation (IBE). The way that Roche and Sober must conceive of IBE for their argument to apply is revealed by their recent (2019) response to Lange’s (2017) objections to their screening-off argument. Roche and Sober’s argument against IBE requires that under IBE, the “explanatory virtues” possessed by a hypothesis H make a difference to a rational agent’s confidence in H independent of that agent’s background opinions. It is unfair to attribute to IBE such a magical connection of explanatory considerations to confirmation.

KEYWORDS: *Bayesianism, Evidential Relevance, Explanatoriness, Inference to the Best Explanation, Sober, Screening Off, Unification*

I. INTRODUCTION

Whether “Inference to the Best Explanation” (IBE) is rational and compatible with a broadly probabilistic account of confirmation have long been subjects of disagreement. (Notable papers on this topic in-

clude van Fraassen (1989), Lipton (2001), and Salmon (2001a), (2001b)). Roche and Sober (2013) gave a new argument against IBE, which has received some critical attention [e.g., McCain and Poston (2014), Climenhaga (2017), and Lange (2017)]. Recently Roche and Sober (2019) have significantly elaborated their argument and clarified how it is supposed to undermine IBE. My aim in this paper is to respond to these new arguments.

My main result is to reveal precisely what makes the versions of IBE targeted by Roche and Sober (hereafter “R&S”) vulnerable to R&S’s refutation. I will argue that R&S presuppose that under IBE, the explanatory virtues that a hypothesis H would possess (if H were true) must have their impact on a rational agent’s degree of confidence in H *independent* of the agent’s background opinions. I will argue that this is an implausible picture of IBE. It depicts explanatory virtues as having an impact on confirmation *magically* rather than thanks to the same sorts of background opinions that determine the confirmatory impact of every other consideration. Properly understood, I argue, IBE is not vulnerable to R&S’s argument.

The best way to arrive at this improved understanding of IBE – and to contrast it with the way that R&S construe IBE – is to follow the same path as R&S (2019) take: by beginning with R&S’s (2019) responses to Lange’s (2017) objections to R&S’s original (2013) argument. Whereas R&S’s original argument purports to give a scientifically realistic case where (R&S argue) explanatory considerations are rendered confirmatorily irrelevant (“screened off”), Lange’s reply purports to give several scientifically realistic examples where explanatory considerations are not screened off. In section II, after briefly rehearsing R&S’s (2013) argument, I will reply to R&S’s objection to Lange’s first example and in section III, I will reply to R&S’s objection to another of Lange’s examples. The lessons thereby learned will enable me to identify (in section IV) the key aspect of R&S’s conception of IBE and what makes it so implausible.

II. ROCHE AND SOBER’S SCREENING-OFF ARGUMENT AND LANGE’S ROBBERY EXAMPLE

Here is R&S’s [(2013), pp. 660-661] key example where explanatory considerations are confirmatorily irrelevant. Let H be that Joe was a heavy smoker before age 50 and let O 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 knows nothing more about Joe. The agent should use her sample frequency information to guide her degrees of confidence to $\Pr(H) \approx .3$ and $\Pr(H | O) \approx .7$. R&S then maintain that the further discovery of *EXPL* (that if *H* and *O* obtained, then *H* would explain *O*) has no confirmatory impact on *H*: $\Pr(H | O) = \Pr(H | O \& EXPL)$. That is, the frequency data underwrite the estimate 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). According to R&S, *O*’s screening off *EXPL* from *H* nicely exemplifies the evidential irrelevance of explanatory considerations.

Lange begins his reply by proposing a purportedly realistic case where *O* fails to screen off *EXPL* from *H*:

... suppose that *H* is that Jones is the person who stole the jewel from the safe, *O* is that the single strand of hair found inside the safe was blond, and the background information tells us that there was exactly one robber and one strand of hair found inside the safe, that Jones has blond hair, and that such a hair has a serious (though not overwhelming) likelihood to have been left by the robber during the robbery (though there are other ways in which the hair could have gotten into the safe). The background information also tells us that Jones is a serious suspect, unlike many other people with blond hair — although Jones is one among several serious suspects with blond hair and there is also a fair likelihood that the robber is not listed among our serious suspects. Background also tells us that if the hair were Jones’s, then Jones would probably be the robber (since he would have left it during the robbery); Jones would have had no occasion to access the safe except to rob it.

Accordingly, since the hair that was found is the same colour as Jones’s hair, *O* lends some support to *H* — $\Pr(H | O) > \Pr(H)$ — though this support is less than maximal, considering that the hair may not have come from the robber and that, even if it did come from the robber, the robber need not be Jones since many other people (including some other serious suspects) have blond hair [Lange (2017), p. 305].

Lange [pp. 305-7] then considers *EXPL*: if Jones were the robber and the single strand of hair found inside the safe were blond, then that Jones is the robber would explain why the strand of hair found in the safe is blond. Lange argues that $\Pr(H | O \& EXPL) > \Pr(H | O)$, i.e., that *O* fails to screen off *EXPL* from *H*. Lange’s argument is that *EXPL*

diminishes one of the concerns weakening O 's degree of support for H , namely, doubt about whether the hair comes from the robber.

R&S (2019) reply that *EXPL* eliminates a possibility: that Jones is the robber but the strand of hair found in the safe doesn't belong to him because it wasn't left by the robber. However, R&S argue, it would be "strange" [p. 124] for the elimination of this possibility to count in H 's favor since this

is a possibility in which H is true. How is it that by eliminating a possibility in which H is true, *EXPL* increases H 's probability? Instead, why not think that by eliminating a possibility in which H is true, *EXPL* decreases H 's probability? [p. 124].

R&S recognize that quirky background opinions could be cooked up under which an observation "raise[s] the probability of a hypothesis and also eliminates a possibility in which the hypothesis is true" [p. 124], namely, where the observation "also eliminates various possibilities in which H is false" [p. 125]. But no such thing is happening in Lange's case, "assuming with Lange, that the background information on hand is realistic" [p. 125].

It seems to me that contrary to R&S, no artificial background opinions are needed to motivate Lange's conclusion that $\Pr(H \mid O \ \& \ EXPL) > \Pr(H \mid O)$. Realistically, that the strand of hair found in the safe was left behind by the robber, if the robber were Jones, confirms to some degree that the strand of hair found in the safe was left by the robber, if the robber were somebody else. That facts about Jones would explain facts about the strand of hair found inside the safe, if Jones were the robber, confirms that if Smith were the robber, then facts about Smith would explain facts about the strand of hair found inside the safe.¹ *EXPL* enhances O 's support for H because *EXPL* traces the hair back to the robber — conclusively so, if the robber is Jones, but to some degree even if the robber is not Jones.

Thus, *EXPL* realistically does more than merely eliminate a possibility in which H is true: that Jones is the robber, but the strand of hair found in the safe doesn't belong to him because it wasn't left by the robber [p. 124]. Rather, *EXPL* realistically diminishes the plausibility of various possibilities in which H is false, such as that Smith (already known to have brown hair) is the robber but the strand of hair found in the safe doesn't belong to him because it wasn't left by the robber. Realistically, *EXPL* strengthens the plausibility of an explanatory connection between

the robber and the hair, thereby making the fact that the hair is blond stronger evidence that the robber is blond — confirming (more strongly than *O* alone does) that Jones is the robber and disconfirming (more strongly than *O* alone does) that Smith is the robber.

Of course, *EXPL* does not *eliminate* the alternatives to *H* that *EXPL* (given *O*) disconfirms. As I mentioned, R&S admit that if *EXPL* eliminated various possibilities in which *H* is false, then *EXPL* might thereby have the overall effect of supporting *H* (beyond the support *O* confers) despite eliminating one way for *H* to be true. But R&S should not have limited themselves to the *elimination* of alternatives to *H*. R&S should also recognize that *EXPL* could support *H* by diminishing the plausibility of (while failing to rule out) various possibilities in which *H* is false.

In short, *EXPL* realistically supports *H* by bearing on the sort of explanation that facts about the hair have: *EXPL* confirms that those facts are explained by facts about the robber. Doubtless we could formulate some (perhaps rather contrived) background opinions under which *EXPL* (despite establishing that the strand of hair found in the safe belongs to the robber, if the robber is Jones and the hair is blond) fails to confirm that the strand of hair found in the safe belongs to the robber if the robber is *not* Jones. But what would that show? A defender of IBE should think that when a hypothesis's explanatory potential² influences its confirmation, that influence is mediated by various empirically justified background opinions. That is far more plausible than that a hypothesis's explanatory potential affects its confirmation *independent* of such background opinions — that is, entirely by virtue of background opinions that are somehow rationally compulsory and *a priori* justified. Philosophers have long found it difficult to see how it could be rationally compulsory and *a priori* justified to hold background opinions according to which hypotheses displaying more explanatory virtues, such as unification and parsimony, always start out more likely than hypotheses displaying less of these virtues.³ Like any other consideration, *EXPL* always (on my view) has its confirmatory impact only by the grace of empirical background opinions. (I will return to R&S's conception of IBE in section IV.)

III. ROCHE AND SOBER'S REPLY TO LANGE'S PHYSICS EXAMPLE

Lange [pp. 308-11] purports to give another example, drawn from the history of physics, where *O* fails to screen off *H* from *EXPL*.

Whereas Lange sees this example as illustrating a scientifically important way in which screening off fails, R&S argue that his example fails.

The idea behind Lange's example is that oftentimes in science, if an agent believes that some live hypothesis H (if true) would be a common explainer of two phenomena (if they obtained), then this belief enables one of the hypothetical phenomena to confirm the other where otherwise one would not have confirmed the other. That is, if the agent were to rule H out or were to believe that H , even if true, would not be a common explainer of the two phenomena, then the agent would not have any reason to take one phenomenon (if it occurred) as confirming the other.

Lange illustrates this idea with two phenomena (the blackbody spectrum and the photoelectric effect) both predicted by H : that the light-quantum hypothesis is empirically adequate. (Let O be that the blackbody spectrum is as if light is quantized and let O^* be that the photoelectric effect is as if light is quantized.) In the early twentieth century, a few physicists (notably Einstein) took H not only as a live hypothesis, but also as perhaps able to explain the two phenomena if they obtained and H were true; that is, some scientists took *EXPL* as holding. They did so because they took it as a serious epistemic possibility that if the light-quantum hypothesis were empirically adequate, then this would be no coincidence; rather, it would be because light is indeed quantized. Those scientists regarded the black body spectrum as confirming H and thereby confirming H 's other bold predictions, such as an equation for the photoelectric effect. By contrast, most physicists took H as a live hypothesis but believed that if it were true, then that various phenomena are as if light is quantized would be just a coincidence (because light would not in fact be quantized). Those scientists denied *EXPL*; they held that were H true, then that various phenomena are as if light is quantized would not help to explain those phenomena. Rather, these phenomena are explained by some other theory of the interaction of light and matter (some non-quantum theory — presumably some theory that treats light as waves). This theory's truth — and that the equations it entails for the blackbody spectrum and the photoelectric effect are coincidentally the same equations as the light-quantum hypothesis entails — explains why these phenomena are as if the light-quantum hypothesis holds. These scientists regarded the blackbody spectrum as confirming H , but not as thereby confirming H 's other bold predictions (such as the equation for photoelectric effect) because it would be just a fluke (a coincidence) for the two phenomena both to be as if light is quantized.⁴ That is, these scientists regarded the blackbody spectrum as confirming

H , but as doing so less strongly than Einstein's allies took the blackbody spectrum to confirm H , since only Einstein's allies took the blackbody spectrum to confirm H 's predictions regarding other phenomena.

In this historical case (according to Lange), the degree to which Einstein's allies regarded H as confirmed by O ($\Pr(H \mid O \ \& \ EXPL)$) exceeded the degree to which the physicists who took H to be coincidental if true (and so took O as unable to confirm H 's bold predictions regarding other phenomena, such as O^*) regarded H as confirmed by O ($\Pr(H \mid O)$). In other words, Lange maintains that in this historical case, $\Pr(H \mid O \ \& \ EXPL) > \Pr(H \mid O)$, i.e., that O fails to screen off $EXPL$ from H , because $\Pr(O^* \mid O \ \& \ EXPL) > \Pr(O^* \mid O \ \& \sim EXPL)$. Lange depicts this as a highly typical case where scientists' regard for a given theory's explanatory potential makes one phenomenon O that it would explain confirmatorily relevant to another phenomenon O^* that it would explain. The scientists' background belief in H 's explanatory potential therefore renders H better confirmed by O than H would be without that belief, since in the latter case, O would fail to confirm H 's implication O^* .

R&S object:

It's not immediately obvious, however, how inequalities like [$\Pr(O^* \mid O \ \& \ EXPL) \gg \Pr(O^* \mid O \ \& \sim EXPL)$] figure in [Lange's] overall argument. The problem is that [the above inequality] makes no mention of H , whereas the claim that $\Pr(H \mid O \ \& \ EXPL) > \Pr(H \mid O)$ does [p. 127].

This seems undermotivated to me. O^* is the proposition that the photoelectric effect is as if there were light quanta, i.e., as if H obtained. O is the proposition that the blackbody spectrum is as if H obtained. So although H does not actually appear in the inequality $\Pr(O^* \mid O \ \& \ EXPL) \gg \Pr(O^* \mid O \ \& \sim EXPL)$, H is not too far away (since O and O^* figure in the inequality). It is not unrealistic (indeed, what happened in the early twentieth century was) for a physicist to regard O (the blackbody spectrum's according with H) as confirming H more strongly insofar as she regarded O as confirming more strongly H 's implications for other phenomena (such as the photoelectric effect) concerning light's interaction with matter.

This sort of thing happens all the time. Let $HEADS$ be that the next three tosses of this coin will land heads and let F be that the first toss lands heads. F confirms $HEADS$ at least by eliminating the possibility that the first of the three tosses lands tails. But F might additionally confirm $HEADS$ by confirming to some degree that the next two tosses land heads. (Whether F does so will depend on the agent's background

knowledge — e.g., whether the agent already knows that the coin is fair.) F will confirm *HEADS* more strongly insofar as F confirms more strongly *HEADS*'s implications for the other two tosses. Likewise, in Lange's example, O eliminates the possibility that the blackbody spectrum will differ from H 's prediction, but O may do more than this to confirm H . O will confirm H more strongly insofar as O not only eliminates one way for H to be false, but also confirms more strongly H 's implications for other phenomena.

R&S press their objection above by noting that Lange could get from $\Pr(O^* \mid O \& EXPL) \gg \Pr(O^* \mid O \& \sim EXPL)$ to $\Pr(H \mid O \& EXPL) > \Pr(H \mid O)$ by using Hempel's "Converse Consequence Condition" (CCC) for incremental confirmation, which says (for any propositions X , Z , and Z^*) that if X confirms Z by some increment and Z^* entails Z , then X confirms Z^* by some increment. R&S point out that from $\Pr(O^* \mid O \& EXPL) > \Pr(O^* \mid O \& \sim EXPL)$, it follows that *EXPL* (given background belief O) incrementally confirms O^* , and since H (that the light-quantum hypothesis is empirically adequate) entails O^* (that the photoelectric effect, in particular, is as if there were light quanta), the CCC would underwrite concluding that *EXPL* (given background belief O) incrementally confirms H , i.e., that $\Pr(H \mid O \& EXPL) > \Pr(H \mid O)$. But of course, R&S intend this suggestion that CCC underwrites Lange's argument as a poison pill: "if this is how the argument is supposed to work, then the argument fails. As is well known. CCC has counterexamples" [p.127].

Lange should refuse to use CCC to reach $\Pr(H \mid O \& EXPL) > \Pr(H \mid O)$ from $\Pr(O^* \mid O \& EXPL) \gg \Pr(O^* \mid O \& \sim EXPL)$. As we have seen, he has no need to appeal to CCC; it is not the only way to support his argument. In the above coin example as well as in the light-quantum example, realistic background opinions enable the hypothesis to be more strongly confirmed by the success of one of its predictions insofar as that success also confirms certain of its other predictions. This need not be the case for all possible background opinions in order for it to hold of certain realistic background opinions. Lange argues that since many scientifically important cases work in the same way as the light-quantum case, there are many such cases in which the evidence fails to screen off the hypothesis from explanatory considerations.

Should an agent regard O^* (that the photoelectric effect's equation is as if light is quantized) as more likely given O (that the blackbody spectrum's equation is as if light is quantized) and *EXPL* than given $O \& \sim EXPL$? That is, should an agent's opinions satisfy $\Pr(O^* \mid O \& EXPL) > \Pr(O^* \mid$

$O \& \sim EXPL$)? The answer obviously depends on the agent's background opinions. Although *EXPL* depicts the two phenomena as having a common explanation in the light-quantum hypothesis, common explainers are not automatically to be deemed more likely than separate explainers; it does not follow somehow solely from the principles of rationality that *O* bears more strongly on O^* given *EXPL* than given $\sim EXPL$. Rather, early twentieth century physicists needed to have some good reason to believe that with one phenomenon having been discovered to be as if there were light quanta, a second phenomenon so being is more likely when the light quantum hypothesis would explain them both than when it would be just a fluke that they are as if light is quantized.

But it is not difficult to see how some early twentieth century physicists had such a good reason. We all have extensive past experience with cases where two phenomena were found to be alike in various unexpected respects (such as cases where two students turned in papers with word-for-word identical paragraphs or cases where two individuals reported having made the same, generally rare observation). In those cases, it turned out more often than it generally does that there is a common reason for the two phenomena. For instance, presumably in our experience, the frequency of one student's paper having a certain long string of words, given that another student's paper also includes that string, is greater among pairs of students whose work had a proximate common cause (such as collaboration, independent use of the same source, or plagiarism of one student by the other) than among pairs of students whose work had no proximate common cause so that any similarity between their work was just a fluke. (Of course, this inequality is a contingent fact, not one that we are entitled to believe solely on the basis of the principles of rationality.) On the basis of this sort of experience, physicists were entitled to $\Pr(O^* \mid O \& EXPL) > \Pr(O^* \mid O \& \sim EXPL)$.

Physicists were so entitled even though they obviously had no prior experience of any pair of phenomena being as if light is quantized and turning out to be explained by light being quantized. Physicists' derived their entitlement from their past experience with phenomena far removed from (and far less esoteric than) blackbodies and photoelectrons. Common, ordinary phenomena have similarities and those similarities are often discovered to have (or not to have) common explainers. Furthermore, in the light-quantum case, there were many other relevant pieces of background knowledge, such as the enormous empirical success of the wave theory of light (which counted against taking *EXPL* se-

riously). These further factors led to the differences of opinion among physicists about whether to take *EXPL* as a live hypothesis. Perhaps both the vast majority of physicists (who did not initially take *EXPL* seriously) and Einstein and his allies (who did) were initially entitled to their opinions about *EXPL*. Presumably, there was greater agreement among physicists about whether (in view of their past experience) two students' papers having a common proximate cause increases the likelihood that the two papers share a long string of words. The role of background opinions about the kind of explanations that certain phenomena are liable to have (based on past experience of the explanations that other cases have) is crucial, in my view, to the way that explanatory considerations such as *EXPL* bear on confirmation. (I will return to this point in the next section.)

In short, that \sim *EXPL* depicts it as coincidental for the two phenomena both to be as if light is quantized, whereas *EXPL* depicts it as no fluke, does not suffice to require an agent to regard *EXPL* as increasing *O*'s bearing on *O**. As R&S say [p. 129], "everything depends on the background assumptions that pertain." But Lange can accept this, too, while maintaining that among early twentieth century physicists, regarding *EXPL* as a serious epistemic possibility was justified and those physicists who had greater confidence in *EXPL* (such as Einstein and his allies) were thereby justified in taking *O* as stronger evidence for *O** – and that similar things happen in many scientifically important cases. In those cases, background opinions do not enable the evidence to screen off the hypothesis from explanatory considerations such as *EXPL*.

IV. AGAINST ROCHE AND SOBER'S CONCEPTION OF IBE

Where, then, do matters now stand between R&S and Lange? Lange has not disputed R&S's original argument that in R&S's smoking example, frequency data render explanatory considerations confirmatorily irrelevant. As we have seen, R&S do not accept Lange's examples as cases where explanatorily considerations are not screened off by the observations *O*. Yet R&S recognize that there exist realistic cases where the background does not allow *O* to screen off *EXPL* from *H*. Nevertheless, R&S hold that this is compatible with their contention that "there are many realistic cases" [p. 130] in which the background opinions do allow the observations to screen off *EXPL* from *H*. This is something that Lange can also accept.

So far, then, there seems little for R&S and Lange to disagree about. Their crucial disagreement concerns the correct lessons regarding IBE to be drawn from all of the above. R&S argue that although explanatory considerations are not *always* screened off, that they are *sometimes* screened off is strong enough to undermine many formulations of IBE (including a “relatively standard formulation” [p. 131] and thousands of others [p. 140n.17]). All of the formulations of IBE that R&S consider are rules specifying that it is rational to believe (or to put greater confidence in) a hypothesis *H* over any of its rivals in the wake of observation *O*, under certain conditions. The conditions required by the various IBE rules, though differing in their details, all demand that *H* possess some significant explanatory advantages with regard to *O*. These advantages may include that *H* would (if true) explain *O* better than any of its rivals would, that *H*'s potential explanation of *O* exhibit a high quantity of the various “explanatory virtues” (which is what makes *H*'s potential explanation of *O* so good), that this high quantity be much higher than any of *H*'s rivals possess, and so forth. (These various advantages are presumably intended by R&S as various ways of cashing out the notion of “best explanation” in IBE.) R&S are then in a position to refute any such IBE rule by the following argument, which they deploy repeatedly [pp. 132, 135, 137]. R&S have already shown that there are realistic cases where explanatory considerations are screened off from *H* by background information and *O*. From this result, they infer that it cannot be that in *every* case within an IBE rule's scope, the rule is correct in deeming the hypothesis possessing these explanatory advantages to be better supported by the evidence. Any such rule is thereby refuted. To refute any such IBE rule, R&S do not require that explanatory considerations be screened off in *all* such cases falling under the rule's scope — merely in *some* of them.

In my view, the problem with this argument against IBE is that IBE is badly represented by the kind of rule that this argument targets. The targeted kind of rule requires that explanatory considerations make their impact on confirmation *independent* of background opinions. That is, these rules depict explanatory considerations as having the same sort of impact whatever the background opinions are. This is what makes these rules vulnerable to R&S's argument, since R&S show that under some background opinions, explanatory considerations have no impact. But this depiction of explanatory considerations as having their confirmatory impact independent of background opinions is what makes these rules

poor representations of IBE. As we saw in both sections II and III, it is precisely the agent's background opinions about what would explain some possible observation that enable some hypothesis to gain greater confirmation from that observation because of the hypothesis' explanatory potential: its potential to give that observation the kind of explanation that the agent believes it is likely to have.

As I mentioned above, the IBE rules that R&S target are rules specifying that it is rational to believe (or to put greater confidence in) a hypothesis H over any of its rivals, in the wake of observation O , if H possesses some combination of "explanatory virtues" with regard to O . R&S [pp. 132-3] list various virtues that (they say) IBE-ers commonly cite: empirical adequacy, explanatory power, fit with background data, fertility, internal consistency, internal coherence, mechanism, parsimony, precision, scope, and unification. But this list makes their conception of IBE a straw man. As I said, the targeted IBE rules require that a hypothesis's possession of various explanatory virtues never be screened off, but instead still favor the hypothesis given whatever background. If $SCORE$ specifies how well or badly H "scores" [p. 136] overall on the explanatory virtues, then the targeted IBE rules all require that $\Pr(H|O \& SCORE)$ always be unequal to $\Pr(H|O)$ rather than that $SCORE$ sometimes be screened off from H by O . But for many of these alleged explanatory virtues, H 's score on these explanatory virtues will already be reflected in $\Pr(H|O)$ without any need to conditionalize on $SCORE$; H 's score will be screened off. Obviously, if H is internally inconsistent, then $\Pr(H|O) = \Pr(H|O \& SCORE) = 0$. Similarly, H 's empirical adequacy and fit with background data will be reflected in $\Pr(H|O)$ already since $\Pr(H|O)$ is implicitly $\Pr(H|O \& \text{background})$. Likewise, if (along the lines of section III) our background opinions (on the basis of our prior experience with explanations of other phenomena that we deem similar to O) lead us to think that O 's explanation is likely to involve a mechanism or a parsimonious batch of causes or laws, then H 's specifying a mechanism or exhibiting parsimony will count in H 's favor, but these contributions to H 's plausibility will already be reflected in $\Pr(H|O)$; further conditionalization on H 's score in these respects will make no difference. Because the IBE rules that R&S target fail to capture the way that explanatory considerations actually enter into confirmation, R&S's refutation of these rules fails to impugn IBE.

The main problem with these rules is that they require explanatory considerations to acquire their confirmatory significance *magically*. Every other kind of evidence acquires its confirmatory impact by way of back-

ground opinions. But these IBE rules require explanatory considerations to make their impact independent of background opinions — as if rationality alone were enough to make a hypothesis's parsimony or the unification it posits or its appeal to a mechanism or its fitting nicely with certain other claims (which we happen to believe already) count in its favor. It is unfair to attribute to IBE such a *magical* connection of explanatory considerations to confirmation. Explanatory considerations must acquire their confirmatory significance through background opinions in exactly the same way as any other consideration does.⁵

Regarding explanatory considerations (such as that *H*, if true, would give a parsimonious, mechanistic, unified, and otherwise allegedly highly virtuous explanation of *O*, if true), suppose I am correct in maintaining that these considerations operate (in enhancing *O*'s confirmation of *H*) no differently from any other considerations (in that *H*'s explanatory virtues count in *H*'s favor only given background confidence that *O*'s explanation, if *O* obtains, is parsimonious, mechanistic, unified, etc.). A critic of IBE might then ask [paraphrasing R&S on p. 131]: Why give any special prominence in a theory of confirmation to *H*'s being *O*'s “best (potential) explanation”?

The answer is that background opinions about the kind of explanation that some known fact is likely to have are an important sort of background opinion. Agents judging among rival hypotheses are often entitled to hold such opinions even (though not solely or always) in cases where they have very little other relevant information, such as where they have no information about frequencies of the kind figuring in R&S's smoking example. (The light-quantum case is a good example of a case where the agents had relatively little relevant background information but where they were warranted in holding opinions about explanations. Although the wave theory of light had been well confirmed, it does not concern the interaction of light with matter, which is central to the blackbody spectrum, the photoelectric effect, and the other phenomena that ultimately told in favor of the light-quantum hypothesis. Very little was previously understood about light's interactions with matter.) We begin our epistemic lives knowing a great many explanations of various things we have observed. Therefore, we have lots of experience with explanations upon which to draw when we consider new hypotheses. We are often in a position to judge whether some new hypothesis would supply various known facts with the kinds of explanations that we justly

expect those facts to have. The explanatory potential of those hypotheses thus turns out to be influential in evaluating them.

The apparent contrast between seeing IBE and seeing background knowledge as responsible for the confirmatory significance of H 's explanatory potential is a false contrast. IBE should be understood as holding that certain features of H 's potential explanations often tell in favor (or against) H by virtue of background knowledge of *other explanations*.

This approach to IBE does not associate it with any list of special “explanatory virtues.” Our prior experience with the explanations of various phenomena that we deem likely to be similar explanationwise to O could lead us to expect O to have a disunified, complicated explanation involving various coincidences — if those are the kinds of explanations that these other phenomena tend to have. In that case, IBE would tend to promote our assigning greater credence to hypotheses that would, if true, provide O with explanations having these “unparsimonious” features. Only magical thinking about IBE would suggest that IBE requires there to be a fixed list of “explanatory virtues” the possession of which by any hypothesis would make it more plausible, whatever our background knowledge may be.

R&S's argument against IBE presupposes that IBE is useful in understanding confirmation only if there is a fixed list of “explanatory virtues” that possess their confirmatory significance magically. Fans of IBE should resist magical thinking.

Department of Philosophy
The University of North Carolina at Chapel Hill
 101B Caldwell Hall
 240 East Cameron Av.,
 Chapel Hill, NC 27599-3125 USA
 E-mail: mlange@email.unc.edu

NOTES

¹ Analogy: It is easy to imagine realistic background opinions under which the discovery that were you to press the button, then your doing so would cause an explosion confirms (by some increment) that were I to press the button, then my doing so would explain an explosion.

² I agree with R&S that “the claim that H is a potential explanation of O is in effect *EXPL*” [p. 131].

³ Sober (2015) tells some of this story.

⁴ Analogy: If we already know (for certain) that a given coin is fair, then the hypothesis that it lands heads on each of its next three tosses is confirmed

(rising from $1/8$ to $1/4$ credence) by the first toss landing heads. But the hypothesis's predictions regarding the second and third tosses are not confirmed; each remains at $1/2$ credence. This occurs because we already know that if the hypothesis were true, then it would be just a fluke (a coincidence) that the three tosses all land heads, since the coin is fair. In this analogy, that the first toss lands heads, the second toss lands heads, and so forth are analogous to the blackbody spectrum being as if light is quantized, the photoelectric effect being as if light is quantized, and so forth. I will return to this analogy shortly.

⁵ Of course, I cannot come close to giving an adequate argument here for this general claim. One objection that might be made is that some "explanatory virtues" acquire their confirmatory impact independent of background opinions. For instance, Harman (1965), p. 89, mentions "plausibility" as contributing to explanatory quality, and as Salmon (2001b), p. 125, remarks, "I would take plausibility to be the very same thing as prior probability." No background opinions are needed to enable a hypothesis's prior probability to have an impact on its posterior probability. However, as Lipton (2001), pp. 93-4, 105, emphasizes, explanatory quality (what Lipton calls "loveliness") must be distinguished from likeliness, since IBE would shed no light on confirmation if it were the view that a hypothesis's likeliness enters into our justification for our degree of confidence in a given hypothesis. So we should not take a hypothesis's "plausibility" as an explanatory consideration that, according to IBE, sometimes bears on our degree of justified confidence in the hypothesis.

One way to construe an "explanatory virtue" so that it has its confirmatory impact independent of background opinions is for the requisite background opinions to be built directly into the virtue itself. For instance, Schubach and Sprenger (2011) have proposed that $\frac{pr(h|e) - pr(h|\sim e)}{pr(h|e) + pr(h|\sim e)}$ measures how well the potential explanation of e supplied by h (when there is one) would explain e . Schubach (2017), p. 48, shows that b_1 's measure exceeds b_2 's iff $pr(e|b_1) > pr(e|b_2)$. In this way, rather than the "explanatory virtue" being defined without reference to background opinions but requiring suitable background opinions to make it confirmatorily relevant, these philosophers have defined the "explanatory virtue" in terms of background opinions in such a way that whenever the "virtue" is present, it is confirmatorily relevant.

Whether this particular account of "loveliness" fits our pretheoretic intuitions about explanatory quality and fits scientific practice (and whether such accounts can be given for particular explanatory virtues rather than for loveliness all things considered) are not my concerns here. My point is merely that although this sort of strategy makes the explanatory quality out to have confirmatory relevance independent of background opinions, this strategy does so only because it bakes the requisite background opinions into what it calls the explanatory quality. This strategy, then, merely shifts the place where background opinions make

a difference to the confirmatory impact of explanatory considerations. It does not reveal that there is no such impact.

REFERENCES

- CLIMENHAGA, N. (2017), “How Explanation Guides Confirmation”; *Philosophy of Science*, 84, pp. 359-368.
- HARMAN, G. (1965), “Inference to the Best Explanation”; *Philosophical Review*, 75, pp. 88-95.
- LANGE, M. (2017), “The Evidential Relevance of Explanatoriness: A Reply to Roche and Sober”; *Analysis*, 77, pp. 303-312.
- LIPTON, P. (2001), “Is Explanation A Guide to Inference?”; in G. Hon and S. Rakover (eds.), *Explanation: Theoretical Approaches and Applications*, Dordrecht: Kluwer, pp. 93-120.
- MCCAIN, K., and T. POSTON (2014), “Why Explanatoriness is Evidentially Relevant”; *Thought*, 3, pp. 145-153.
- ROCHE, W., and E. SOBER (2013), “Explanatoriness is Evidentially Irrelevant, or Inference to the Best Explanation Meets Bayesian Confirmation Theory”; *Analysis*, 73, pp. 659-668.
- (2019), “Inference to the Best Explanation and the Screening-Off Challenge”; **teorema**, 38, pp. 121-142.
- SALMON, W. (2001a), “Explanation and Confirmation: A Bayesian Critique of Inference to the Best Explanation”; in G. Hon and S. Rakover (eds.), *Explanation: Theoretical Approaches and Applications*, Dordrecht: Kluwer, pp. 61-92.
- SALMON, W. (2001b), “Reflections of a Bashful Bayesian”; in G. Hon and S. Rakover (eds.), *Explanation: Theoretical Approaches and Applications*, Dordrecht: Kluwer, pp. 121-136.
- SCHUPBACH, J. (2017), “Inference to the Best Explanation, Cleaned Up and Made Respectable”; in K. McCain and T. Poston (eds.), *Best Explanations: New Essays on Inference to the Best Explanation*, Oxford: Oxford University Press, pp. 39-61.
- SCHUPBACH, J., and J. SPRENGER (2011), “The Logic of Explanatory Power”; *Philosophy of Science*, 78, pp. 105-127.
- SOBER, E. (2015), *Ockham's Razors – A User's Manual*; Cambridge: Cambridge University Press.
- VAN FRAASSEN, B. (1989), *Laws and Symmetry*; Oxford: Oxford University Press.